

# TRUE Is False and Why It Matters

Robert William Fischer

Texas State University-San Marcos

## 1. Introduction

The proponents of inference to the best explanation (IBE) are willing to move from a judgment about the quality of an explanation to a judgment about its probability. In other words, they sanction inferences that have this form:

- P1: Facts  $f_1 - f_n$  obtain.
- P2: If true, hypothesis  $h$  would offer a better explanation of  $f_1 - f_n$  than would any competing hypothesis.
- C: So, probably,  $h$  is true.<sup>1</sup>

This formulation raises an obvious question: for a given hypothesis, in what sense is it alleged to offer a ‘better’ explanation of  $f_1 - f_n$ ? The standard answer is that the hypothesis has a higher score on the explanatory virtues: conservatism, modesty, simplicity, generality, and predictive power.<sup>2</sup> But this answer appears to be problematic. Here Bas van Fraassen’s objection to it:

Judgments of simplicity and explanatory power are the intuitive and natural vehicle for expressing our epistemic appraisal. [But these] are specifically human concerns, a function of our interests and pleasures, which make some theories more valuable or appealing to us than others. Values of this sort [...] provide reasons for using a theory, or contemplating it, whether or not we think it true, and cannot rationally guide our epistemic attitudes and decisions. For example, if it matters more to us to have one sort of question answered rather than another, that is no reason to think that a theory which answers more of the first sort of questions is more likely to be true (not even with the

---

<sup>1</sup> Depending on your views about explanation, this argument may need an additional premise: something to the effect of, “if true, hypothesis  $h$  would provide a satisfactory explanation of  $f_1 - f_n$ .” However, nothing here turns on its inclusion.

<sup>2</sup> This particular list is due to Quine and Ullian, with ‘generality’ substituted for the more awkward ‘fecundity’ (Quine and Ullian 1978, 64-82). It is not unusual. For very similar ones, see (Lycan 1988, 130) and (Lipton 2004, 122). Obviously, not everyone characterizes IBE this way. For example, at best, Bayesians construe IBE as a heuristic tool for fixing the priors and likelihoods. The debate with the Bayesian is an important one, but I can set it aside here: I am taking for granted the conception of IBE that my interlocutors are taking for granted (at least with respect to the objection that I discuss in the main text).

proviso ‘everything else being equal’). It is merely a reason to prefer that theory in another respect (Van Fraassen 1980, 87).<sup>3</sup>

IBE faces a slew of objections, many of which are formidable. However, let’s bracket all but the one that appears above. Does it provide a *distinct* problem for those who regard IBE as a source of epistemic justification? In other words, is IBE faulty simply because it relies on “specifically human concerns” that are “a function of our interests and pleasures?” No.

Let’s get clearer about van Fraassen’s argument. It seems to go as follows:

- P1: If reason *r* is an epistemic reason for (subject *s* to believe) *p*, then *r* increases the likelihood of *p*’s truth.
- P2: But IBE’s reason(s) for (subject *s* to believe) hypothesis *h* do not increase the likelihood of *h*’s truth.
- C: Therefore, the reasons given by inference to the best explanation are not epistemic reasons.

Let’s call this *the argument from the truth-conduciveness of epistemic reasons* (ATER). If ATER is sound – and if (plausibly enough) you need epistemic reasons to get epistemic justification – then it seems that IBE cannot provide us with epistemic justification.

Some respond to ATER by attacking P2.<sup>4</sup> However, I tend to think that P2 is true. My quarrel is with P1: it is *not* the case that all epistemic reasons increase the likelihood of truth.<sup>5</sup> The claim that they do assumes a form of epistemic value monism – which, I’ll argue, even IBE’s critics should reject. In my view, then, the objection above amounts to the observation that

---

<sup>3</sup> We find the same argument in one of Scott Shalkowski’s recent papers: “Reasons are sometimes epistemic, sometimes pragmatic. IBE is proposed as a general kind of inference involving epistemic reasons; it is to provide us with reasons to adopt a theory as more likely to be true than its competitors and not merely as a tool useful for accomplishing some non-alethic goal. [...] Simplicity is a theoretical virtue, let us grant, but it is an instrumental virtue. Simple theories are easier to work with, so recognizing that a theory is simple provides one with a reason to work with the theory, but this is a conclusion of a piece of practical reasoning” (Shalkowski 2010, 171-172).

<sup>4</sup> Richard Swinburne, for example, contends that “it is a fundamental *a priori* principle” that simpler theories are more likely to be true than are more complex ones (Swinburne 2001, 102). And there are a number of less radical defenses of simplicity: e.g., (Quine and Ullian 1978), (Sober 1981), and (Kelly 2007).

<sup>5</sup> There are, of course, philosophers who insist that the individual virtues are individually truth-conducive. Simplicity is usually taken to be the hardest one to defend, but Richard Swinburne nevertheless contends that “it is a fundamental *a priori* principle” that simpler theories are more likely to be true than are more complex ones (Swinburne 2001, 102). There are also a number of less radical defenses of simplicity – e.g., (Quine and Ullian 1978, 69-70), (Sober 1981, 145), and (Kelly 2007, 561).

IBE is incompatible with epistemic value monism, and that is no objection at all.

## 2. From ATER to TRUE

We need to begin by distinguishing two ways of interpreting ATER's first premise:

P1: If reason  $r$  is an epistemic reason for (subject  $s$  to believe)  $p$ , then  $r$  increases the likelihood of  $p$ 's truth.

On the flat-footed reading, P1 imposes a necessary *external* condition on epistemic reasons: namely, that they must increase the likelihood of truth. But if this is the correct reading, then there are two strikes against van Fraassen's relying on P1. First, no proponent of IBE needs each explanatory virtue to be *individually* truth-conducive; IBE does not require that, for example, simpler theories are more likely to be true just in virtue of their simplicity. Rather, the proponent of IBE needs it to be the case that the virtues are *jointly* truth-conducive. There is no obvious reason why various non-truth-conducive virtues might not 'cancel one another out', allowing the reasoner to triangulate the truth, as it were.<sup>6</sup> Second, and more importantly, while epistemic reasons may need to satisfy an external condition, it's hard to see how making this point would help van Fraassen. To assess IBE's reliability, we would need to check whether (a) there is a positive correlation between those propositions supported by the reason in question and those propositions that are true and (b) a negative correlation between those propositions supported by the reason in question and those propositions that are false. But propositions don't wear their truth values on their sleeves, so we can only go on our best judgments. And as soon as we admit this, we must also recognize that there will be disagreement: if I think that our judgments about the existence and properties of unobservables are generally accurate, then I will be inclined to say that reasons supporting those judgments are truth-conducive; if van Fraassen doesn't, then he

---

<sup>6</sup> For more on this point, see (Fischer, ms).

won't be so inclined. So on the 'necessary external condition' reading of P1, the merits of ATER turn on the merits of P2 (which alleges that IBE's reason(s) for hypothesis *h* do not increase the likelihood of *h*'s truth). And our judgments about the merits of P2 will depend on two factors: first, the list of hypotheses that we believe to be justifiable via IBE; and second, the list of those hypotheses that we take to be true. But van Fraassen is using this argument (among many others) to *motivate* shortening the list of hypotheses that we take to be true; i.e., he is trying to argue that IBE cannot justify beliefs about unobservables. Hence, the 'necessary external condition' reading of P1 does not help his project; it seems to beg the question at hand. What is the alternative?

I propose that P1 concerns the *aim* of epistemic reasons. We might reformulate ATER accordingly:

- P1\*: If reason *r* is an epistemic reason for (subject *s* to believe) *p*, then *r* is aimed at increasing the likelihood of *p*'s truth.
- P2\*: But IBE's reason(s) for (subject *s* to believe) hypothesis *h* are not aimed at increasing the likelihood of *h*'s truth.
- C: Therefore, the reasons given by inference to the best explanation are not epistemic reasons.

I think that this reading fits more naturally with the passage quoted above; at any rate, it avoids the problems just mentioned. It also make it clear why, earlier, I posited a connection between P1 and epistemic value monism. P1\* insists that the only epistemically valuable feature of a reason is its being aimed at truth. Hence, P1\* commits its proponent to a version of epistemic value monism – the view that reasons have only one epistemically relevant feature.

### 3. TRUE

What's wrong with P1\*? It will be easier to see this if we take a detour through ethical theory.<sup>7</sup>

---

<sup>7</sup> My argument in this and the next section is inspired by (Lycan 1988, Chapter 7). I do not mean to suggest that

Utilitarianism, at least in its simple, hedonistic form, is committed to both *value monism* and *proceduralism*. An ethical theory is committed to the former if it maintains that all situations have only one morally relevant feature; according to utilitarianism, that feature is well-being. An ethical theory is committed to the latter just in case it says that there is a decision procedure for determining whether an action is obligatory, permissible, or wrong; for utilitarians, this is the principle that you should maximize well-being.<sup>8</sup> With these two points in mind, and idealizing a bit, we can represent utilitarianism with a function: it takes a set of action / outcome pairs as inputs, selects the one with the greatest overall well-being, and gives the action that leads to that situation as the output; that action, of course, is the one that utilitarians judge to be obligatory.<sup>9</sup>

The function just outlined represents act utilitarianism. How would we need to modify it in order to represent rule utilitarianism? We replace the set of action / outcome pairs with a set of slightly more complex pairs, the first member of which is a candidate moral rule, the second of which is the outcome that would result from universal adherence to that rule. The function still selects the one with the greatest well-being. However, instead of giving an obligatory action as an output, it gives a moral rule; we then apply the rule to our situation to determine what's obligatory.

It's easy to reframe rule utilitarianism as an epistemological position. Instead of candidate moral rules, the first member of each pair is a candidate epistemic policy; instead of global outcomes, the second member of each pair is the number of truths that would be believed if that policy were followed.<sup>10</sup> Instead of selecting the outcome with the greatest well-being, our

---

Lycan would agree with anything that I say here.

<sup>8</sup> I'm using the phrase 'decision procedure' loosely, where it doesn't imply that informed and competent agents are always in a position to carry it out.

<sup>9</sup> Here is one respect in which this is an idealized representation: like Stalnaker's account of counterfactuals, it assumes that there are no ties.

<sup>10</sup> I am treating wellbeing as a simple property; hence, the parallel with truths believed. Later, I'll discuss a variant that balances truths believed with falsehoods avoided.

new function selects the one with the greatest number of resultant true beliefs, giving that epistemic policy as an output. As before, the output is not itself the obligatory action; rather, the output is the principle that determines what you ought to believe in a given circumstance – equivalently (so say I), it determines the belief that you would be justified in holding in those circumstances. Let's call this view *truth-maximizing rule utilitarianism in epistemology* (TRUE). Like its cousin, TRUE is a version of value monism: it takes truth to be the only feature of a belief that is of worth. Also like its cousin, TRUE is a form of proceduralism: it takes there to be a straightforward decision procedure that settles which of the many possible epistemic policies is correct. Why does it recommend maximizing true beliefs? As in ethics, your theory of value drives your theory of the right: if you think that only well-being is of moral worth, then it is hard to see what you would recommend *other* than maximizing well-being. After all, if well-being is of moral value, then surely more is better, at least if all other things are equal. And if value monism is true, then all other things always *are* equal – there is never anything else with which well-being competes. So, you should maximize it. The same argument applies, *mutatis mutandis*, to truth given TRUE.

TRUE is probably not just a form of epistemic value monism: it is probably the only epistemological position that is plausible if epistemic value monism is correct. As I suggested in the preceding paragraph, it's likely that epistemic value monists are committed to an epistemology that is structurally analogous to utilitarianism. But in epistemology, the analog of act utilitarianism is hopeless: *that* view would say that a belief is justified iff it's true, since (a) such a view would only take into consideration the local features of the belief and (b) such a view would take the truth of that belief to be the only feature that matters. But, of course, it is not the case that a belief is justified iff it's true. The analog of rule utilitarianism, TRUE, avoids this

problem by introducing the epistemic policies: they are designed to take non-local factors into account – namely, the number of true beliefs that would be achieved given universal adherence to the epistemic policy – thereby preventing TRUE from having the awkward consequence that sinks the epistemic analog of act utilitarianism.

#### 4. We Should Reject TRUE

However, as sane as it may sound, TRUE has very implausible implications. Here is the argument. I suggested that we can represent TRUE as a function: the inputs are policy / success rate pairs, the output is the most truth-conducive epistemic policy. I also intimated that ‘being the most truth-conducive epistemic policy’ means ‘being the policy that produces the greatest number of true beliefs if it were followed’. But this can’t be right. The policy that will do best here is the one that tells us to believe *everything*. If truth is the *only* valuable doxastic feature, then there is no value to avoiding falsehood. So, if we were to believe every proposition and its negation, then we wouldn’t miss out on any truths, thereby maximizing what’s of epistemic value.<sup>11</sup> But this is ridiculous.

To avoid this problem, we should make a friendly amendment to TRUE. We’ll still say that truth is still the only valuable doxastic feature, but we’ll add a principle called ‘NOFALSITY’, according to which believing falsely has epistemic *disvalue*. Call our revised version of TRUE – i.e., the conjunction of TRUE and NOFALSITY – ‘T&~F’. T&~F preserves the spirit of TRUE, if not the letter. Problem solved?

No. Now, the most straightforward interpretation of ‘being the most truth-conducive

---

<sup>11</sup> *Objection:* We can’t believe contradictions, epistemic policies create epistemic obligations, and we aren’t obligated to do the impossible; so, we can’t be obligated to believe every proposition and its negation, which means that this policy is not in the running. *Reply:* It’s not at all clear to me that we can’t believe explicit contradictions. But even if that’s right, then we certainly *can* believe implicit contradictions. In other words, even if we can’t believe  $p$  &  $\sim p$ , it’s surely the case that we can believe  $p$  and we can believe  $\sim p$ .

epistemic policy' is something like 'being the policy that maximizes the ratio of truths to falsehoods believed'. This looks like a recipe for radical epistemic caution: if you take this policy seriously, then you should believe only self-evident truths. If you believe even one falsehood, then it doesn't matter how many truths you believe, since your ratio of true to false beliefs will invariably be lower than it would have been had you believed no falsehoods at all. But as long as you find at least one self-evident truth (the *cogito* or your favorite tautology) and you believe no falsehoods whatever, your ratio will be as high as it possibly can be.<sup>12</sup>

So we seem to be torn between two extremes: either radical epistemic abandon (believe everything) or radical epistemic caution (believe only the self-evident). You might object I'm assuming both more and less control over our beliefs than is plausible, ignoring:

- (a) that you can't believe whatever you want, so you can't believe everything (which is supposed to undermine my objection to TRUE), and
- (b) that so many of our beliefs form spontaneously, so we can't limit ourselves to believing a single self-evident truth (which is supposed to undermine my objection to T&~F).

I grant both (a) and (b), but they make no trouble for my argument. Concerning (a), is it really just your *inability* to believe everything that makes it a terrible epistemic policy? If TRUE is correct, then this seems to be the case. Surely it isn't, though. Even if it *were* psychologically possible to believe indiscriminately, that would not be a way of securing justified beliefs. And the same point applies to (b): even if it *were* psychologically possible to believe only the self-

---

<sup>12</sup> *Objection:* Any number over zero isn't a ratio (it's ill-defined); so, you would have to believe at least one falsehood to achieve the goal of maximizing the *ratio* of true to false beliefs. *Reply:* First, the 'maximize the ratio' formulation isn't mine; it's common enough in the literature: see, e.g., (Nozick 1993, 69). Second, it's easy enough to recast the conversation in terms of maximizing the percentage of truths believed, in which case my argument stands. And third, you can still make the ratio version work. Suppose that you believe one self-evident truth and believe its negation; you then believe as many propositions as you can that are logically equivalent to the self-evident truth. Since there are infinitely many of them, you can make the ratio as high as your mental capacities permit (and this with minimal epistemic risk). *Objection:* Logically equivalent propositions are equivalent, period; so, this solution puts your ratio at .5. *Reply:* Logically equivalent propositions are not equivalent, period. If they were, then 'red is a color' and '2 + 2 = 4' would express the same proposition, since they both express necessary truths. And that's absurd.



evident, would this be a good epistemic policy? If  $T \& \sim F$  is correct, then the answer is ‘Yes’. But surely this would be a mistake.

Here is a further consideration. Perhaps some beliefs are inescapable: even if we judge them to be false, we cannot abandon them. If there are such beliefs, though, and we indeed judge them to be false, then surely we can still recognize the epistemic tension that this creates. I suspect that something similar is the case when we judge the risk of error to be unacceptably high: whether or not we can actually abandon the beliefs in question, if we judge the risk of error to be too great, then surely we can judge them to be epistemically subpar. But when is the risk excessive? If TRUE is correct, then our only advice is to believe as many truths as possible; it follows that the risk is *never* excessive. If a belief is epistemically subpar just in case the risk of being wrong crosses some threshold, TRUE seems to suggest that we should *never* judge a belief to be epistemically subpar. Alternately, if  $T \& \sim F$  is correct, then our only advice is to maximize the ratio of truths to falsehoods believed; now, the risk is excessive whenever there is a threat that we might *not* maximize that ratio, which is to say that it’s excessive whenever we believe what isn’t self-evident.  $T \& \sim F$  seems to suggest, then, that we should almost *always* judge our beliefs to be epistemically subpar. So, whether supplemented with NOFALSITY or not, TRUE is in trouble.

Someone might object that it’s uncharitable to articulate either TRUE or  $T \& \sim F$  in terms of truth *simpliciter*. Rather, they should be cashed out in terms of *significant* truths (i.e., “maximize the number of significant truths believed’ or ‘maximize the ratio of significant truths to falsehoods believed’). I agree that it should be, but the proponents of TRUE and  $T \& \sim F$  cannot. What makes some truths significant while others are not? Whatever it is, it’s something other than their mere truth – perhaps their usefulness, or their explanatory power, or their fit with

what we believed pre-theoretically, or what have you. And crucially, either the significance of a belief is explicable solely in terms of its truth, or it isn't. In other words, significance is either extrinsically or intrinsically valuable. If it's extrinsically valuable, then significance won't save either TRUE or T&~F from the problems that I've been detailing, since there will never be a case in which significance trumps truth, thereby giving you a reason to take an epistemic risk. But if significance *isn't* explicable solely in terms of its truth – i.e., if it's intrinsically valuable – then to set significance alongside truth is to reject epistemic value monism, and hence to reject TRUE and T&~F.

## **5. Back to IBE**

I grant that I may have overlooked a perfectly good policy that's based on the assumption that truth is the only thing of epistemic worth; if so, then TRUE's devotees should provide it. Suppose they can't. How should we diagnose the problem? Well, as I've indicated, rationally increasing your stock of beliefs beyond the self-evident requires a policy about the management of epistemic risk. Whatever policy you adopt, it will need to give advice having the following form: *risk error only if...*, where the ellipses stand for something *else* of epistemic worth. Your policy might be, for example, that you should risk error only if the proposition would increase the coherence of your belief system. Alternately, you may maintain that you should risk error only if the proposition in question seems to be true, absent any defeaters – this is the way taken by those who defend 'phenomenal conservatism' (e.g., (Huemer 2001)). If you go the first route, then you're assigning epistemic value to coherence; if you go the second, then you're assigning it to conservatism. There are no doubt plenty of other options, but they'll all lead you to assign intrinsic epistemic value to something other than truth. In other words, they'll lead you to deny

epistemic value monism. But now recall ATER:

- P1\*: If reason  $r$  is an epistemic reason for (subject  $s$  to believe)  $p$ , then  $r$  is aimed at increasing the likelihood of  $p$ 's truth.
- P2\*: But IBE's reason(s) for (subject  $s$  to believe) hypothesis  $h$  are not aimed at increasing the likelihood of  $h$ 's truth.
- C: Therefore, the reasons given by inference to the best explanation are not epistemic reasons.

If epistemic value monism is false, then P1\* is false. So P1\* is false. IBE may face a number of serious challenges, but ATER is not among them.

## References

- Kelly, Kevin. 2007. A New Solution to the Puzzle of Simplicity. *Philosophy of Science* 74:561-73.
- Lipton, Peter. 2004. *Inference to the Best Explanation*. International Library of Philosophy. New York: Routledge.
- Lycan, William G. 1988. *Judgement and Justification*. New York: Cambridge University Press.
- Quine, W. V. and J. S. Ullian. 1978. *The Web of Belief*. New York: Random House.
- Shalkowski, Scott. 2010. IBE, GMR, and Metaphysical Projects. Pages 167-187 in *Modality : Metaphysics, Logic, and Epistemology*. Edited by Bob Hale and Aviv Hoffmann. New York: Oxford University Press.
- Sober, Elliott. 1981. The Principle of Parsimony. *British Journal for the Philosophy of Science* 32:145-56.
- Swinburne, Richard. 2001. *Epistemic Justification*. New York: Oxford University Press.
- Van Fraassen, Bas C. 1980. *The Scientific Image*. New York: Oxford University Press.