PROBABILITY AND DEFEATERS

BY

ALVIN PLANTINGA

Abstract: Branden Fitelson and Elliott Sober raise several objections to my evolutionary argument against naturalism; I reply to four of them.

My thanks to Branden Fitelson and Elliott Sober (hereafter 'F&S') for their comments on my evolutionary argument against naturalism.[1] F&S devote most of their attention to what I called "the preliminary argument" (WPF, pp. 228–29). This argument as stated in WPF contains an error: it confuses the unconditional objective or logical probability of R with its probability conditional on our background knowledge.[2] The main argument, happily, is unaffected, and here I'll comment only on what F&S have to say about the main argument. F&S start several hares, most of which seem to me to run rather badly. I can't chase them all, so I'll restrict myself to the following four.

1. P(R/N&E) low or inscrutable?

Let P be the proposition *my cognitive faculties are reliable*, N be philosophical naturalism (i.e., the proposition that there is no such person as God or anything like God) and E the proposition that we and our cognitive faculties have come to be by way of the processes favored in contemporary evolutionary theory. In WPF I argue that P(R/N&E) is low or inscrutable, and do so by considering several mutually exclusive and jointly exhaustive subcases P_i. These 4 subcases are ways in which belief can be related to behavior: (1) epiphenomenalism (belief isn't causally connected with behavior at all), (2) semantic epiphenomenalism (belief is causally connected with behavior, but just by virtue of its *neurophysiological* properties and not by virtue of its *content*), (3) belief is causally

related to behavior by virtue of content as well as neurophysiological properties, but is maladaptive, and (4) the common sense possibility (call it 'C') according to which belief is both adaptive and also causally efficacious by way of content as well as neurophysiological properties. I argued that $P(R/N\&E\&P_i)$ is low or inscrutable on all but C; I also said that $P(R/N\&E\&C)$ is not very high. The next step of the argument was to try to arrive at a value for $P(R/N\&E)$, given the values for $P(R/N\&E\&P_i)$ for each of the P_i. This proceeds in terms of a weighted average principle: $P(R/N\&E)$ is equal to the weighted average of $P(R/N\&E\&P_i)$, weighted in each case by $P(P_i/N\&E)$. The probability calculus gives us a formula here:

$$P(R/N\&E) = (P(R/N\&E\&P_1) \times P(P_1/N\&E)) + (P(R/N\&E\&P_2) \\ \times P(P_2/N\&E)) + (P(R/N\&E\&P_3) \times P(P_3/N\&E)) \\ + (P(R/N\&E\&P_4) \times P(P_4/N\&E)).$$

I argued that $P(R/N\&E\&C)$ is not very high (or inscrutable) and that $P(R/N\&E\&P_i)$ for each of the other P_i is low (or inscrutable); the result is that $P(R/N\&E)$ is itself low or inscrutable.

F&S don't quarrel with my suggestions here, except in the case of $P(R/N\&E\&C)$. There they mistakenly represent me (p. 121) as arguing that that probability is low; what I actually said, however, is that "Here one does not quite know what to say about the probability that their cognitive systems would produce mainly true beliefs, but perhaps it would be reasonable to estimate it as somewhat greater than 1/2" (WPF p. 227). Then later I amend this by adding that suggestion that it is also reasonable to take this probability to be inscrutable. More important, F&S consider *only* $P(R/N\&E\&C)$. But given the structure of the argument, it might be that $P(R/N\&E)$ is low or inscrutable even if $P(R/N\&E\&C)$ is very high – indeed, as high as you please. So even if their animadversions on my argument for at best a moderately high value for $P(R/N\&E\&C)$ were on target, the overall argument wouldn't suffer.

But they aren't on target. My argument is inconclusive, say F&S, because, so they say, I neglect to consider the fact that whether a trait is likely to evolve depends upon the *availability* of the trait:

Plantinga's mistake here is that he ignores the fact that the probability of a trait's evolving depends not just on its fitness, but on its *availability.* The reason zebras don't have machine guns with which to repel lion attacks is not that firing machine guns would have been less adaptive than simply running away; the trait didn't evolve because it was not available as a variation on which selection could act ancestrally. . . . By ignoring the question of availability, Plantinga, in effect, assumes that natural selection acts on the set of *conceivable* variants. (pp. 120–21).

Well, it's nice to know that machine guns weren't originally available to zebras, although even if they were it wouldn't have helped much – their

hooves are much too big and clumsy to operate those little triggers. But I entirely fail to see how my argument in effect assumes that natural selection acts on the set of *conceivable* variants – e.g., the conceivable variant in which creatures from an advanced civilization elsewhere in the galaxy take pity on our Australopithicene ancestors struggling in epistemic darkness and by a bit of genetic reengineering confer reliable cognitive faculties on them. What EAAN assumes is only that the epistemic probability of false beliefs (or protobeliefs, or whatever would be the appropriate precursors of belief) being available to natural selection is not much lower than that of true beliefs being available. And that seems clearly plausible: the natural thing to think is that any time at which beliefs are available is a time at which both true and false beliefs are available. On the other hand, one who disputes this assumption will presumably think P(R/N&E&C) is inscrutable, which is also in accord with EAAN.

2.   *Is* (N&E) & P(R/N&E) is low or inscrutable *a defeater for R?*

F&S say no: "After all, people who believe *E&N* might have other reasons for believing *R*. For example, they might argue that *R* is a basic proposition that does not need theoretical support, or that *R* derives its epistemic credentials from something other than the thesis of evolutionary naturalism" (p. 124). They might indeed argue for both those theses, and if they did, I think they'd be right. But that fact does not insulate R from defeat for the naturalist. My belief that I see a sheep is a basic proposition that does not require theoretical support; nevertheless, it is possible to get a defeater for it (as when you, the local authority on sheep, point out that what I see is only a sheep dog that looks like a sheep from this distance). So it is entirely possible to acquire a defeater for a belief you hold in the basic way. In the same way, R (we may suppose) is a basic proposition not needing theoretical support: nevertheless the naturalist's belief *(N&E) & P(R/N&E) is low or inscrutable* can perfectly well provide him with a defeater for it. Similarly, the fact that R does not derive its epistemic credentials from N&E does not deliver it from defeat. For example, I believe I've taken XX, a substance I believe sometimes induces massive unreliability; I also believe that P(R/I've taken XX) is low or inscrutable. I then have a defeater for R with respect to myself; but of course R does not derive its epistemic credentials from my belief that I've taken XX. So it is entirely possible to acquire a defeater D for a belief, even if the belief does not derive its epistemic credentials from D. And just this is the case, if EAAN is right, with respect to R. R has its warrant in the basic way, and does not derive it from N&E; nevertheless *(N&E) & P(R/N&E) is low or inscrutable* provides the naturalist with a Humean defeater[3] for R.

F&S also suggest that:

> People who believe *E&N* should not regard the fact that this probability is "inscrutable" to them as a reason to reject *R*. We suspect that many people who are well acquainted with the theory of special relativity and who think that birds fly still don't know what value to assign to Pr(Special relativity/birds fly).

Here the point seems to be that:

> (1) For any propositions *A* and *B*, if P(*B*/*A*) is inscrutable, then *A* is a defeater for *B*

is false. Indeed it is; but of course my argument doesn't for a moment presuppose that it is true.[4] There are cases like this where *A* *is* a defeater for *B*, and cases where it is *not*; this case, so I say, is one of the former.

F&S seem also to think that ignorance of probabilities is never a guide to belief: "The Principle of Indifference is flawed because it claims to obtain probabilities from ignorance; the start of Plantinga's main argument makes the complementary mistake of holding that ignorance of probabilities is a guide to belief" (p. 124). But surely ignorance sometimes *is* a guide to belief. In the trivial case, your ignorance of P(*A*/*B*) is a good reason for not believing, e.g., that P(*A*/*B*) is .23. But there are many less trivial cases. I assume, as usual, that the thermometer T I've just bought is reliable. You then tell me that this thermometer was made in a factory F whose Luddite owner aims to do his best to frustrate modern industrial society by fabricating instruments many of which are unreliable, but you don't know the ratio of reliable to unreliable instruments produced by the factory. P(T is reliable/T was fabricated in F) is then inscrutable for you, as it will be for me if I believe you. This constitutes a defeater, for me, of my initial assumption that T is reliable, and it is a defeater, in part, because of ignorance of a probability.

3.   *Conditionalization*

"Notice that Plantinga assumes that evolutionary naturalists have no basis for deciding what to think about *R*, other than the proposition *E&N* itself. This crucial assumption is never defended in either *Warrant and Proper Function* or 'Naturalism Defeated'" (p. 125). Now the fact is I think R has *intrinsic* warrant, warrant in the basic way. That is because I believe human beings have been made in God's image, part of which involves our being able to form true beliefs and acquire knowledge; and the module of the design plan governing the production of the belief that

R is aimed at the production of true belief. So of course I don't think that the naturalist's only basis for what to think about R is N&E. (But I do think that neither the naturalist nor anyone else can give a cogent *argument* for R.) So how to understand F&S here? Perhaps the best way to understand this complaint is as an attempt to raise the conditionalization problem: what, in this context, can the naturalist properly conditionalize on? Just N&E itself? Presumably not: but then what further? This is indeed a tough problem, but not, so far as I can see, in such a way as to give the naturalist an escape route.[5]

4.   *A defeater for all of my beliefs?*

F&S also suggest that even if I have a defeater for R, I don't necessarily have a defeater for the bulk of my beliefs: "Even if *E&N* defeats the claim that 'at least 90% of our beliefs are true', it does not follow that *E&N* also defeats the more modest claim that 'at least 50% of our beliefs are true'. Plantinga must show that *E&N* not only defeats *R*, but also defeats the claim that 'at least a non-negligible minority of our beliefs are true' (p. 125). But why must I show that? I agree that *(E&N) & P(R/N&E) is low or inscrutable* does not offer a direct defeater for the belief that at least 50% of our beliefs are true. But how does that help the naturalist? F&S seem to make the mistake, here, of thinking that if you don't have a defeater for the proposition that 50% of your beliefs are true, then 50% of your beliefs are such that you don't have a defeater for them. But that doesn't follow at all. I still have a defeater for each of my beliefs, even if I also believe that 50% of them are true, and even if 50% of them *are* true. I've been reading an authoritative book on evolutionary biology, naturally enough believing what I read. The author then tells me that up to 50% of the propositions written on p. 45 of the book are false, but does not tell me which. If I have no further relevant information I will not believe any of these propositions, even though I believe that at least 50% of them are true. Another example: I've been keeping records, relying on that thermometer T of a few paragraphs ago. You then tell me that the factory owner designed T in such a way that 50% of its readings are true and 50% false. If I believe you and have no other source of relevant information, I have a defeater for each of the beliefs I formed on the basis of the readings of the thermometer, even though I believe half of them are true.

   Finally, I'd like to comment on F&S's paragraph 2.3, p. 125. This paragraph illustrates what seems to me a couple of unhappy features of their essay: (1) consistent misunderstanding, and (2) attempts to show that my arguments are uncogent by finding invalid argument forms of which they are instances, or false propositions that entail their premises (or both). But *every* argument is an instance of an invalid argument form (e.g.,

$P_1, \ldots, P_n$; therefore Q) and every proposition is a consequence of some false proposition.[6] In para 2.3 they begin by saying that they have "interpreted the main argument . . . as asserting that a low value for Pr($R/E\&N$) suffices for $E\&N$ to defeat R", which is near enough true. But then they say that in "Naturalism Defeated" "Plantinga denies that this is what he meant and tries to develop an account of defeat that clarifies how the argument is supposed to go." I don't know what they mean here. I do think (near enough) that a low value for Pr(R/E&N) suffices for E&N to defeat R; but I certainly did not (in "Naturalism Defeated" or anywhere else) deny that this is what I meant. So what do they have in mind? I *think* perhaps they are referring to my rejecting (in "Naturalism Defeated") the idea that

(2) For *any A* and *B*, if P($R/A\&B$) is low or inscrutable, then *B* is a defeater for *A*.

But I never so much as dreamt that *that* proposition is true, and certainly did not propose or presuppose it in my original statement of EAAN.

Next, F&S note that I develop three principles of defeat in "Naturalism Defeated" and complain that "he never explains how they are relevant to establishing that *E&N* defeats *R*. In fact their logical form renders them incapable of closing the gap between premises 1 and 2 [their statement of my argument on p. 123]." Now to close the gap between those premises, I take it, would be to show that if,

(3) P(R/N&E) is low,

then,

(4) E&N is a defeater for R.

But of course those three principles were not designed to show that (3) implies (4). Thus consider the First Principle of Defeat:

(FPD) If *S* rationally believes that the warrant a belief *B* has for him is derivative from the warrant a belief *A* has for him, then *B* is not a defeater, for him, of *A* ("Naturalism Defeated", p. 48).

(FPD) was designed to summarize a number of cases in which (2), above, is false: those cases where the warrant of *B* is derivative, in the appropriate way, from the warrant of *A*. Similarly, the Second Principle of Defeat is designed to show some other circumstances under which a belief *B* does not constitute a defeater for a belief *A*; and the third principle is designed to exhibit a set of circumstances under which a proposition *B* does not

constitute a defeater-defeater for a proposition *A*. Hence the complaint that they fail to fill the gap between premises (1) and (2) is based just on misunderstanding.

Next, F&S propose that (FPD), far from filling the gap between (3) and (4) "apparently helps establish that *E&N* is *not* a defeater of *R*"; that is because they "suspect that many evolutionary naturalists rationally believe that their warrant for believing *E&N* depends on their being warranted in believing that their cognitive faculties are highly reliable". But it is one thing for the warrant *B* has for me to *depend on A*'s having warrant for me, and quite another for the warrant *B* has for me to be *derivative from* the warrant *A* has for me. As I said in "Naturalism Defeated", the paradigm case of the latter is where I *infer B* from *A*. But the naturalist does not, one hopes, infer N&E from R. I don't know what, if anything, he does infer N from, but presumably he infers E (perhaps by way of an "argument to best explanation") from premises about the fossil record, the distribution of species across the world, certain results from molecular biology, and the like. This argument certainly does not require the additional premise that his cognitive faculties are reliable, and will not be the least improved by its addition.

They next correctly point out that X's being a defeater of Y does not depend upon P(Y/X) being low, at any rate if *self-defeat* is possible; but again, I never for a moment thought otherwise. They go on to claim that it is difficult to connect low or inscrutable probability with defeat. As we have seen, however, for each there are plenty of unproblematic cases where it is indeed connected with defeat. Finally, they suggest (p. 126) that I waffle on the proposition *high probability is necessary for rational belief* and suggest that "What Plantinga is coming up against here is a close relative of the phenomenon that Kyburg's lottery paradox made vivid". They then add that "This connection with the lottery paradox suggests that the task of repairing the main argument is formidable". But I can't see that F&S have done anything to show that the main argument *needs* repair. Further, I fail to see any connection between the lottery paradox and my argument. The lottery paradox arises when one asserts quite generally that rational belief that *A* depends upon *A*'s being probable with respect to some body of propositions – what you *know*, perhaps, or what you *know immediately*, or what is *certain* for you, or something of the sort. I certainly don't believe and as far as I know have never suggested that high probability (with respect to what?) is in general necessary for rational belief. Thus I believe I've just thrown three heads in a row with this fair coin; even if that belief is improbable with respect to the rest of what I believe it is nevertheless perfectly rational. Still, further, EAAN, as far as I can see, doesn't even depend on the proposition that naturalists are rational in believing R only if P(R/N&E) is high. A naturalist might rationally believe R in the basic way, for example, and never so much as

consider P(R/N&E); then she will not (or at any rate need not) have a defeater for R. It is only the *reflective* naturalist, one who sees that this probability is low or inscrutable, and thinks about the bearing of this on R, who gets a defeater here. This is another case where F&S find some false general proposition from which what I say perhaps follows, and then complain that what I say is therefore false.[7]

Department of Philosophy
University of Notre Dame

NOTES

[1] *Pacific Philosophical Quarterly* 79 (1998). My argument can be found in *Logos* 12, 1991, pp. 27–50, and in chapter 12 of *Warrant and Proper Function* (New York: Oxford University Press, 1993) (hereafter 'WPF').

[2] I thank Fitelson and Sober for helping me see this. For a correction and repair of the preliminary argument, see my *Warranted Christian Belief* (New York: Oxford University Press, 2000), pp. 229–31.

[3] See my *Warranted Christian Belief* (New York: Oxford University Press, 2000), p. 225.

[4] As I point out in "Naturalism Defeated" (unpublished but available on the web at philofreligion.homestead.com/Plantingapage.html), pp. 46 ff.

[5] See Richard Otte's "Conditional Probabilities in Plantinga's Argument Against Naturalism" in *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument Against Naturalism*, ed. James Beilby (Ithaca: Cornell University Press, 2001), which is largely devoted to the conditionalization problem. I've said what I can about this problem in Section IC of "Reply to Beilby's Minions" in the Beilby volume.

[6] And not just some necessarily false proposition; every possibly true proposition is a consequence of a false contingent proposition.

[7] One more thing: F&S argue that I am mistaken in holding that "if naturalism is true, then, surely, so is evolution" and in holding that the probability of evolution on naturalism is high: "Neither of these claims is right. Recall that proposition E adverts to the mechanisms described in *contemporary* evolutionary theory. If that theory were found wanting, it would not entail the falsehood of naturalism . . ." Quite right: naturalism does not entail evolution. What I meant is that the epistemic probability of evolution is high, given naturalism *together with our current evidence*. That is of course quite compatible with the thought that if new evidence showed up, naturalists could sensibly move to some other theory of evolution, or even give it up altogether.