



A House Divided against Itself Cannot Stand

Plantinga on the Self-Defeat of Evolutionary Naturalism

TIMOTHY O'CONNOR

Alvin Plantinga argues that belief in evolutionary naturalism is self-defeating. Let R denote the thesis that our basic cognitive faculties are mostly reliable, and EN the thesis that human beings and their cognitive faculties arose by means of entirely natural processes of the kind posited by current evolutionary biology, processes unguided and undesigned by God or any other supernatural being. The probability of R on EN, Plantinga plausibly maintains, is inscrutable by us. Since EN is relevant to the truth of R, the inscrutability of R on EN gives the adherent of EN a reason to withhold belief in R: EN is *evidence*, for the naturalist, that calls into question his belief in R. Withholding belief in R clearly would have disastrous implications for one's beliefs, as R's truth underpins the warrant for all our other beliefs. Worse still, Plantinga contends, there is no reasonable means of escaping this predicament once one is mired in it. Since the argument provides a defeater for all the naturalist's beliefs, he is left with nothing that might enable him to defeat the defeater. Hume's game of backgammon beckons.

Since R has a foundational status in our system of beliefs, we can evaluate the force of Plantinga's reasoning only in the context of a general account of epistemic justification or warrant. For the purposes of this essay, I will adopt Plantinga's own 'proper functionalism', on which "a belief has warrant, for a person, if it is produced by her cognitive faculties functioning properly in a

I thank Al Plantinga for helpful correspondence on the subject of this chapter in response to an earlier paper of mine ("An Evolutionary Argument against Naturalism?" *Canadian Journal of Philosophy* 24 [1994]: 527–540). During that time, many a breakfast of mine was given over to reflections such as those below—so much so, that for some time afterward, eating muesli regularly triggered thoughts concerning the machinations of evil demons. I have written the present essay partly in the hope that I may again eat my breakfast in peace.

congenial epistemic environment according to a design plan successfully aimed at [truth]."¹ As with other externalist theories, this account allows that cognitively basic beliefs such as R may be fully warranted despite the fact that a person cannot give a non-question-begging argument on its behalf. It simply need be *true* that our faculties function properly (and aim at truth) in giving rise to this belief—one needn't be in a position to argue that this is (or is likely to be) so.

This, at any rate, is how things initially stand for R, on Plantinga's account. His critique of belief in naturalism presupposes that as we begin to reflect on our beliefs concerning the origins of our cognitive faculties and their probabilistic connections to R, we should begin modifying our confidence in R. In the language of probabilistic reasoning, it is appropriate to 'start off' assigning a very high probability to R, independent of any opinion we may have concerning the absolute a priori probability of R. Subsequently, however, we should *conditionalize* the value we assign to R on factors that are relevant to the probability that R is true. If we believe such factors to obtain and judge the probability of R on them to be inscrutable, we should declare all bets off regarding the truth of R and suspend belief. We should do so because this is precisely how we should proceed in analogous circumstances with more mundane beliefs. Suppose I am driving through the countryside and come to believe there is a sheep in the field. But then I am told by a humorless and reliable friend that this is yet another locale in which the residents like to play tricks on strangers by carefully dressing up dogs to look like sheep. Since I don't know the proportion of sheep to look-alikes in this vicinity, the reasonable course is to refrain from believing that what I see is a real sheep.

Are matters as straightforward as this with respect to R? Consider two examples that Plantinga himself discusses in other contexts. The first is the Case of the Purloined Letter. You are presented with powerful evidence that you have stolen an important letter. You had motive, means, and opportunity. Sadly, it is known that you are not above such shameful activities as filching letters. Finally, there is very strong evidence that you were the only person who could have stolen the letter and it was later recovered from your bureau drawer. Now, before hearing the prosecutor's case, you believed that you had not stolen the letter. And with good reason: you had no reason to doubt your memory in this regard (which was in fact functioning normally), and you distinctly recalled seeing the letter and resisting the temptation to take it. Should you change your belief in the face of the powerful evidence to the contrary? Presumably not. You have countervailing evidence in the form of a distinct memory of not having taken the letter.

Second case. A Christian theist reflects on the dismal facts of pain and suffering in human history. He knows of no convincing account of adequate reasons God might have for permitting these things to occur. He then considers

¹ *Warrant and Proper Function* (New York: Oxford University Press, 1993) 237.

the probability that God exists on this evidence. Suppose that, impressed by our finitude, he judges it unlikely that we could figure out what God's reasons would be for creating a world with such features, were He to exist and have created a world just like ours. He concludes that the conditional probability that God exists on facts about the magnitude and distribution of human suffering is inscrutable. Should he become agnostic? As Plantinga himself emphasizes, if his proper functionalist account of warrant is basically correct, we cannot divorce this question from others concerning God's existence and His intentions concerning the functioning of human cognitive capacities. It is plausible to suppose that if traditional Christian teaching is essentially correct, then God does not intend for human beings to cease believing in Him under such circumstances. On such teaching, a Christian who confidently persists in belief despite not knowing what to say about the evidential bearing of the facts of evil on God's existence is, other things being equal, regulating his beliefs in accordance with God's design plan. It is the tendency toward doubt or disbelief that reflects noetic malfunction.²

These examples highlight the fact that simple conditionalization on new evidence need not always be required within a proper functionalist epistemology. Some beliefs will have a special status, depending on certain facts about the *design plan* of the cognizer in question. An evolutionary naturalist will naturally begin his reflections on Plantinga's challenge by considering the implications EN itself might have for thinking about proper function, just as the theist does in considering the challenge to theism from evil. An astute naturalist will note first of all that the argument doesn't, as a psychological matter, tend to push people to cognitive despair, as Plantinga claims it ought to do. Our belief in R is quite tenacious. Given EN, this is not surprising. There is no plausible evolutionary story one might tell about why our cognitive faculties should be designed to *regulate* this belief strictly in accordance with our evidence for it. Doing so would obviously be disastrous for our coming to true beliefs about other matters. But it would equally undercut our ability to come to a true belief about R itself. It is plausible, given EN, that the naturalist's unshaken confidence in R reflects our cognitive design plan, stemming from a module aimed at truth. We're designed to start off accepting R without evidence and to continue on that way. In view of our cognitive limitations—we're not able to pull off any version of Descartes's project—it is proper that we adopt R and related beliefs as unquestioned framework principles, against which we can sensibly adjudicate our beliefs about less foundational matters.

I think we can say something even stronger. Strictly speaking, there *couldn't* be a defeater for R, for any creatures in any possible world. That would re-

² More needs to be said about the instantiation of a proper functionalist epistemology within a Christian theistic metaphysics to make these claims convincing, of course. There is no better place to begin exploring this matter than Plantinga's own *Warranted Christian Belief* (New York: Oxford University Press, 2000).

quire a *design plan* which reliably aimed at truth in some circumstances, a part of which was that we ought to give up belief in R when we take note of certain of our beliefs about the world and about related conditional probabilities. But as that potential upshot would lead one to abandon the attempt to form true beliefs, it surely could not be part of a design plan with that very end. Or if the concept of a design plan is to be quite broad, then only egregiously flawed instances would contain in this way internal obstacles to their own success.

At this point, Plantinga might protest that our argument has proven too much. Are we not claiming that R is necessarily immune from evidence? And isn't this plainly wrong—couldn't one have, for instance, good reason to think one was or was in the process of becoming insane?

But we are not claiming that R is indefeasible. One could *rationaly* conclude that R is false (in relation to oneself) by having powerful, direct evidence that there *is no design plan successfully aimed at truth*. This would seem reasonable if one's cognitive output were persistently and massively inconsistent—which, I take it, might well capture the circumstances under which one might rationally take oneself to be insane.³ Of course, once one draws this conclusion, it will undercut one's confidence in the very evidence that led to it, thereby giving rise to Hume's loop of reason against itself, at least until one retires to more soothing pursuits such as backgammon. But note that we shouldn't say that one would be *warranted* in drawing this fateful conclusion, for that term of art is tied to the notion of a design plan aimed at truth, and there is no purchase on the idea of such a plan that generates persistent, massive inconsistency. (What then is the operative notion of 'rational' here? It's hard to say precisely—perhaps a basic kind of theoretical rationality which requires a minimal degree of evidential consistency.)

Plantinga, however, argues that massive internal inconsistency is not the only way one might reasonably be led to abandon R. For suppose I come to believe that I am the victim of Descartes's evil demon, who is bent on my being deceived about my true situation. Don't I then have good reason to abandon R?⁴ To evaluate this analogy properly, we must first tighten it up a bit. As the evil demon hypothesis is usually presented, we know that if it's true, then R is false. (By hypothesis, all of my beliefs about the material world are false.) But in the challenge posed to the naturalist, we are urged to say only that the probability of R on evolutionary naturalism is inscrutable. So let us modify the evil demon story accordingly. We will now say that I believe that there is an evil demon who has created my cognitive faculties, but I judge

³ Perhaps a second way is this: rather than believing merely that the general probability of R on a certain broad *type* of origin for our cognitive faculties is low or inscrutable, we come to believe confidently that the specific circumstances of the actual past, together with the evolutionary processes at work, were such that R is improbable on them. But how could we know that, apart from the kind of direct evidence (pervasively inconsistent outputs and the like) discussed in the text?

⁴ "Naturalism Defeated" (unpublished), 51–52.

the probability of R on this origin to be inscrutable. (Perhaps I believe that this demon has created numerous races on various planets, and in at least one case has chosen to massively deceive them, and in at least one other case has chosen to give them reliable faculties. But I have no idea whether either case was an isolated act of whimsy or reflected a decided tendency on his part.)

Once we make this change in our hypothesis, however, it isn't plausible that we are rationally required to abandon R. There isn't massive inconsistency of the sort noted just above, a sort that would make it unreasonable to continue believing there is any coherent design plan underlying one's belief formations. And whether R would have or lack warrant under those circumstances depends on the demon's design plan: has he so designed the cognitive system in question that a properly functioning specimen would abandon R under such circumstances? Well, if the agent is like most naturalists who encounter Plantinga's argument, he is not *inclined* to do so. So the only evidence the agent has in this matter—his own tenacious belief—suggests that he is functioning properly in continuing to believe R. Nothing in the structure of Plantinga's argument as applied to this case would give him reason to think otherwise.

Let me bring the issue into sharper focus by yoking the cases of the Christian theist's response to the evidential bearing of evil on theism and the naturalist's response to the evidential bearing of evolutionary naturalism on R, and then posing a dilemma. Plantinga would say that the probability of theism given the facts of pain and suffering as we know them— $P(T/E)$ —is (at best) inscrutable for the theist. Must the theist, if he is to adjust his beliefs properly, conditionalize on his ignorance, so to speak? There are four possibilities here:

(1) He must conditionalize on his ignorance and thereby come to regard the probability of theism itself as inscrutable for him in the circumstances. If one accepts this, one is also likely to accept Plantinga's own verdict on the case of naturalism. But the larger upshot will be that *every* reflective person should come to withhold belief in R!

(2) He must conditionalize on his ignorance, but he need not come to judge the probability of theism as inscrutable. For he has an independent source of warrant that neutralizes the potential defeater of his belief before it takes hold: the internal testimony of the Holy Spirit that sustains the Christian's confident trust in the existence and character of God. This is a kind of evidence, in a broad sense of the term, that counteracts the potential impact of the inscrutability of $P(T/E)$ for him.⁵ But if the theist should say this, why should not the naturalist say the same concerning the 'evidence' coming from his own tenacious belief that R?

(3) He need not conditionalize on his ignorance or modify his beliefs in

⁵ A brief remark by Plantinga in correspondence several years ago suggested this interpretation to me. But he appears to reject this view of the matter in *Warranted Christian Belief*, 478ff.

any other way. Theism continues to have warrant for him. (But then by parity of reasoning the same should hold for the naturalist.)

(4) He need not conditionalize on his ignorance, but he should adjust his belief concerning $P(T/E)$. If theism is true, it is part of my design plan that I should believe T despite my inability to discern a priori $P(T/E)$. But since it is also part of my design plan that I adjust my beliefs, as best I can, in the direction of greater probabilistic consistency, I should *conclude* that $P(T/E)$ is high. Put differently, I should see on reflection that T 's having a good deal of warrant for me commits me to believing tacitly that the probability of T on all the relevant evidence I possess is high. And because of the special status T has in my God-given design plan, it should trump any beliefs I may hazard concerning conditional probabilities governing T , i.e., my belief in T has more warrant than do any of my beliefs concerning the conditional probability of T on any particular proposition. Although Bayesian accounts of theory confirmation have a difficult time handling this fact, we clearly can rationally revise our estimates of conditional probabilities. (Consider a scientist who comes better to understand the implications of a complex theory and so revises his estimates of the probability of that theory on various bits of potential evidence.) But again, if this is so for the theist, what reason have we to assume it is different for the naturalist? If one is to challenge the propriety of the naturalist's doing this, it cannot depend on assuming facts about the design plan inconsistent with naturalism.

Of these four options, (1) will be implausible to all nonskeptics, and (2) strikes me as stretching the notion of evidence beyond useful limits, so that it becomes little more than a terminological variant on (4). Whichever way we go on (3) or (4)—a verdict that will reflect one's judgment on whether it is appropriate to extend the machinery of probability theory to belief kinematics quite generally—since parallel moves are ready to hand, the naturalist has nothing to fear from Plantinga's argument.