Is Evolutionary Naturalism Epistemologically Self-Defeating?

Paul Churchland

Abstract: Alvin Plantinga argues that our cognitive mechanisms have been selected for their ability to sustain reproductively successful behaviors, not for their ability to track truth. This aspect of our cognitive mechanisms is said to pose a problem for the biological theory of evolution by natural selection in the following way. If our cognitive mechanisms do not provide *any* assurances that the theories generated by them are true, then the fact that evolutionary theory has been generated by them, and even accepted by them, provides no assurance whatever that evolutionary theory is true. Plantinga's argument, I argue, innocently assumes that the (problematic) "truth-tracking character" of our *native* cognitive mechanisms is the *only* possible or available source of rational warrant or justification for evolutionary theory. But it isn't. Plantinga is ignoring the *artificial* mechanisms for theory-creation and theory-evaluation embodied in the complex institutions and procedures of modern science.

My answer to our title question is: Not in the slightest. Indeed, quite the reverse. If anything, evolutionary biology provides us with a partial *explanation* of how it is that cognitive creatures with at least some factual grip on reality can emerge in the first place, develop over the millennia, and eventually penetrate quite deeply into the intricate structures of objective reality. To be sure, nothing guarantees that any creature, or species, will succeed in its representational endeavors. But many creatures will have at least a running chance at successfully representing the world, if only a small part of it. How this can happen, we shall examine shortly.

To begin, however, let us address how our title question arose in the first place. The question arises, first, because *theories* are not the units of natural selection; *genes* are. And second, because natural selection has no direct

Paul Churchland, Philosophy Department, University of California, San Diego

interest in truth *per se*, nor, indeed, any authoritative access to it, even if theories somehow *were* the units of natural selection. Instead, what shapes the meander of biological evolution is the relative *reproductive success* of a creature's *behavior*, not the *truth value* of that creature's *theories*. Accordingly, we can safely infer, from the mere existence of a certain type of creature at a given time, that its genome typically generates individuals whose behavior in its natural environment tends to produce reproductive success, at least relative to the local competition. What we *cannot* infer, from the mere production, existence, and acceptance of a certain *theory* by a certain type of creature, is that this theory is *true*, or even *probably* true. The selective pressures exerted by evolution are focused on quite different factors. Our cognitive mechanisms have been selected for their ability to sustain reproductively successful behaviors, not for their ability to track truth.

This preliminary conclusion, which I accept, is said to pose a problem for the biological theory of evolution by natural selection in the following way. If our cognitive mechanisms do not provide *any* assurances that the theories generated by them are true, then the fact that evolutionary theory has been generated by them, and even accepted by them, provides no assurance whatever that evolutionary theory is true. In sum, if you accept evolutionary theory, then you have undermined the authority of our native cognitive mechanisms to reliably generate true theories, or to discriminate true theories from false ones, and hence, you have undermined our warrant or authority to claim truth for evolutionary theory. Evolutionary theory, according to Prof. Plantinga, is thus self defeating. If it's true, you cannot (apparently) have any significant reason to believe that it's true. Or so the argument goes.[1]

If one shares a fondness for evolutionary theory, and for naturalism in general, one's first dialectical impulse, in response to this second argument, is probably to go back to the first argument (in my second paragraph), dig in one's heels, and attempt to rebuild some sort of bridge between the evolutionary selection of our cognitive mechanisms, on the one hand, and the probable and/or at least the partial truth of the theories that they generate, on the other. This turns out to be a surprisingly awkward undertaking, for a variety of reasons some of which we shall discuss below. Accordingly, the result of pursuing that first impulse is likely to be a fairly quick and disappointed return to Plantinga's original skeptical conclusion against the warrant available for evolutionary naturalism. Later in this paper we shall indeed pursue something like that first bridge-building impulse, and we shall find some interesting results. But that is not what demands our immediate attention.

What compels our attention is a glaring fallacy in Plantinga's second argument. In rejecting evolutionary theory's claim to warranted acceptance of any kind, Plantinga's argument innocently assumes that the (problematic) "truth-tracking character" of our *native* cognitive mechanisms is the *only* possible or available source of rational warrant or justification for evolutionary theory. But it isn't. Plantinga is ignoring the *artificial* mechanisms for theory-creation and theory-evaluation embodied in the complex institu-

tions and procedures of modern science. These super-added mechanisms lie mostly outside the biological brain, and they provide a much more creative environment for generating interesting theories, and a much more demanding filter for evaluating them, than a single biological brain could ever provide with its native resources alone.

Let us remind ourselves of some of the salient elements of this collective artificial mechanism for theoretical cognition. First, the community of the world's scientists is indeed specifically and self-consciously concerned to evaluate the *truth* of the various theories that are severally brought to its attention. The judgments of this highly special community are not based on any indirect criteria concerning the reproductive success of anyone's or anything's typical ecological behavior. Rather, such judgments are based on a wide variety of techniques for testing for the truth of whatever theory is at issue.

These include directly comparing the theory's logical or mathematical predictions with the results of our experimental probings. They include techniques for the evaluation of the statistical significance of our experimental results, for the minimization or elimination of noise, and for the identification of the principal components of variation within large data sets. They include techniques of experimental design and control for avoiding potential confounds. They include techniques for avoiding "cherry-picking" and other forms of experimenter bias, techniques such as 'double blind" procedures to avoid these potential perversions. They include the critical scrutiny of the alleged results by the rest of the professional community and the subsequent attempts to replicate the relevant experimental findings in independent laboratories. They also include evaluating the proposed theory's consistency and explanatory consilience with the already established corpus of (provisional) theoretical wisdom accumulated over centuries of prior scientific activity, similarly collective in nature and similarly technique-driven.

Furthermore, such supra-individual and methodologically sophisticated techniques of rational evaluation are steered by a vast armory of enhanced (artificial) sensory modalities, modalities that reach far beyond the narrow domains revealed by our native sensory equipment. Science employs telescopes (radio, infrared, visible-light, X-ray, and gamma-ray telescopes) to reveal to us the structure and activities of the very large; it uses microscopes (visible light, electron microscopes), and electrochemical techniques such as nucleic-acid sequencers, microelectrodes, and activity-sensitive fluorescent dyes, to reveal to us the structure and activities of the very small; it uses newly sensitive devices such as voltmeters, ammeters, and spectrometers to reveal to us phenomena to which we are congenitally blind for reasons that have nothing to do with spatial scale; and it uses techniques such as radioactive dating, ice-core analysis, and geological stratigraphy for determining structures and activities over deep time.

Such artificial sensory modalities (and these examples merely head a very long list) give us a dramatically enhanced sensory grip on objective

reality, and thus a comparably enhanced *environment* for both the creation and the critical evaluation of new theories. Given the also-enhanced techniques of sheerly rational *evaluation* (of theory in the light of data) scouted two paragraphs ago, the cognitive engine of the Collective Scientific Community can supply rational warrant—both positive and negative—far beyond what can be supplied by a single individual with his native smarts and sensory equipment. In particular, it can supply warrant for the weave of naturalistic claims and explanations that constitute modern evolutionary theory. And it can do so even if we all *agree* that an individual brain "designed" by natural selection has no significant claim to be a "truth tracker". Even so, and nonetheless, an artificial cognitive system—namely, the interlocking set of institutions, procedures, and experimental technologies that make up modern science—a system that has been deliberately "designed" by many generations of scientists to be as reliable a "truth tracker" as we can make it, can indeed supply genuine warrant in favor of a theory such as biological evolution by natural selection. As indeed it has. Even as this theory continues to be fine-tuned and up-dated by unfolding data, Naturalistic Evolutionary Biology must be counted among the very best-confirmed and systematically warranted theories of all of the theories in the pantheon of science, both past and present. Only a handful of other theories even come close.

But singing the praises of Evolutionary Naturalism in particular is a task for a different paper, and perhaps for a different author. My concern here is to point out that the failure of this theory to pick up any automatic warrant from the (admittedly false) presumption of automatic truth tracking on the part of the brains that created it is entirely consistent with its successfully picking up reams of warrant from another source entirely—namely, the deliberately and explicitly truth-tracking mechanisms that constitute the collective institutions of human science. Indeed, warrant from this artificial sociological source carries far *more* rational weight than does any possible warrant from the natural-selection-shaped capacities of the biological brain alone. Accordingly, the absence of the latter sort of warrant leaves Evolutionary Naturalism still safely awash in many dimensions of rational warrant: the very *same* sorts of warrant displayed by all of the other currently accepted theories in science. Plantinga's deflationary argument involves a diversionary fixation on a minor point—indeed, a strictly irrelevant point—a diversion that steers our attention away from the truly important source of warrant for evolutionary naturalism, and for scientific theories in general, namely, the interlocking evaluatory mechanisms of the entire scientific community.

This said, there remains an important issue about the epistemological credentials of our native cognitive mechanisms. And I here repeat my opening agreement with Plantinga's claim that Evolutionary Naturalism, solely on its own, provides little or no reasonable expectation that evolutionarily successful creatures will boast cognitive mechanisms that reliably tend to track Truth in particular. That point unites us, and I would be no less determined to defend it even in the unlikely event that Plantinga were inclined to let it go.

Taken at face value, this justly deflationary estimation of our human cognitive credentials leads us to predict that typical human *theories*—about the origins of mankind, about the structure of the heavens, about the origins of the universe, about the nature of disease, about the causes of motion, and about the nature of life—will be hopelessly parochial, culturally various, and strictly *false*. And so they have been. The compulsive Animism that dominated primitive human cultures; the celebrated Seven Days of Creation at the hands of a Great God embraced by a more recent culture; the Garden-of-Eden account of human origins; the flat, immobile earth enclosed in a Star-flecked Sphere that rotates daily; the Invading-Demon theory of disease; the Eternal Reward/Punishment account of the authority of moral imperatives; the Vital Spirit theory of Living Creatures; all of these, and countless other cognitive embarrassments, typically advanced by and celebrated in the world's popular religions, are just the sorts of benighted stories that you would *expect* of brains originally selected *primarily* for their capacity to engage in reproductively successful behaviors within their enveloping environmental niche. So far then, the predictions of Evolutionary Naturalism are nicely in accord with the (often embarrassing) facts of historical human cognition. There is no conflict with the empirical facts here. Just the reverse.

And yet, there is an upside to the evolutionary story as well, whose outline will serve to bring this essay to a close. I begin by inviting the reader to consider a broader conception of representation, and of *successful* representation, than that embodied in the familiar framework of broadly *sentence*-like representations, and of their *truth*. There are many motives for broadening our conception here, but the most immediately relevant in the present context is that the vast majority of biological creatures throughout the long history of life on Earth have had no capacity whatever for expressing or manipulating representational vehicles even remotely like sentences, and hence no capacity for ever achieving the peculiarly sentential virtue of truth. They have been using other representational schemes entirely, schemes that display dimensions of success and failure quite different from the familiar dichotomy of truth vs. falsity.

Cognitive Neurobiology has already given us an opening grip on what those more primitive, pre-linguaformal schemes of representation consist in, and of how they can embody information about any creature's immediate sensory and practical environment. The suggestion, currently under vigorous development, is that, in response to the ongoing statistical profiles of their complex sensory inputs, nervous systems (even very simple ones) typically develop a high-dimensional *map* of the difference-and-similarity structure of the abstract features typically instanced in and encountered in their sensory environments. The development of such maps is typically achieved by post-natal processes such as Hebbian learning, which has long been known to sculpt any creature's synaptic connections, and hence its acquired internal maps of feature-spaces, in accordance with the fine-grained statistical structures of the creature's sensory inputs.

The take-home point for the present discussion is that the dominant

scheme of representation in biological creatures generally, from the Ordovician to the present, is the internal *map* of a range of possible *types* of sensorily accessible environmental features. Not a sentence, or a system of them, but a *map*. Now a map, of course, achieves its representational successes by displaying some sort of *homomorphism* between its own internal structure and the structure of the objective domain that it purports to portray. And unlike the strictly binary nature of sentential success (a sentence is either truth or it's false), maps can display many different *degrees* of success and failure, and can do so in many distinct dimensions of possible 'faithfulness,' some of which will be relevant to the creature's practical (and reproductive) success, and many of which will not.

The point of this brief excursion into Cognitive Neurobiology is that, if we broaden our conception of representational activity, in biological nervous systems, to embrace the synaptically embodied *feature-space map*, then we *can* find a perspective from which the reproductively-focused selection pressures on the creatures that develop them *will* exert at least an *indirect* pressure in favor of the capacity for generating accurate cognitive maps, for it is precisely those maps that subsequently govern the creature's practical behaviors, including its reproductive behaviors. On the whole, good maps will serve the creature better than will poor ones. Accordingly, Evolutionary Naturalism suggests that there will be a strong tendency for living creatures to develop cognitive feature-maps that are at least *roughly* accurate *partial* portrayals of the *practical* environment in which the creature must make its way. This presumption falls well short of heralding Truth for such representations. But it does serve to explain how pre-human creatures can achieve penetrating internal representations of remarkable intricacy and accuracy, at least on some accountings of accuracy, all within a purely naturalistic universe.[2]

It may also explain how *humans*, too, mostly manage to do it, for the great bulk of human cognition is sublinguaformal as well. And on the negative side, it may also explain why our theories about domains that are far *removed* from our immediate practical experience and control are typically so benighted. The explanation is that, in such domains, our native cognitive mechanisms are plainly "in over their heads." To achieve cognitive success in those more rarefied domains, we need the additional armamentarium of the institutions of modern science, and most especially, their vital means for transcending our native sensory and manipulative limitations. Only then will we have a reasonable chance at cognitive success of any sort, whether accurate maps or true theories.

I close with a lighthearted bit of mischief. The preceding excursion into Cognitive Neurobiology aside, the central point of this paper remains that Evolutionary Naturalism can derive its primary claim to rational warrant from the cognitive operations of the assembled institutions and organs of the worldwide Scientific Community, even though it might find no rational warrant in the unaided cognitive operations of a typical individual human brain. This point finds an ironic parallel with a prominent feature of settled

Catholic dogma, which claims that a person can indeed find Grace & Everlasting Salvation through the continuing administrations of the assembled institutions and organs of God's Holy Church, but cannot reasonably hope to find such Grace & Salvation by operating entirely outside of it. Given the doctrinal sympathies of his home institution (the rightly respected University of Notre Dame), you'd think Prof. Plantinga might have seen that parallel, and taken its message to heart: if Mother Church can achieve such a lofty aim, at least for its more deserving persons, why shouldn't Mother Science achieve the much more modest aim of finding Rational Warrant, at least for its more deserving theories?

NOTES

1. Alvin Plantinga, *Warrant and Proper Function* (New York: Oxford University Press. 1993), Ch. 7.

2. These themes are explored at length in P. M. Churchland, *Plato's Camera: How the Physical Brain Captures a Landscape of Abstract Universals*, a book in progress.