

# Natural Selection Does Care about Truth

Maarten Boudry and Michael Vlerick

*True beliefs are better guides to the world than false ones. This is the common-sense assumption that undergirds theorizing in evolutionary epistemology. According to Alvin Plantinga, however, evolution by natural selection does not care about truth: it cares only about fitness. If our cognitive faculties are the products of blind evolution, we have no reason to trust them, anytime or anywhere. Evolutionary naturalism, consequently, is a self-defeating position. Following up on earlier objections, we uncover three additional flaws in Plantinga's latest formulation of his argument: a failure to appreciate adaptive path dependency, an incoherent conception of content ascription, and a conflation of common-sense and scientific beliefs, which we diagnose as the 'foundationalist fallacy'. More fundamentally, Plantinga's reductive formalism with respect to the issue of cognitive reliability is inadequate to deal with relevant empirical details.*

## 1. Introduction

True beliefs are better guides to the world than false ones. This is the common-sense assumption that undergirds theorizing in evolutionary epistemology. An early hominid who believes that tigers are dangerous predators, and acts upon that belief, will fare better (in terms of survival and reproduction) than one who believes that they are docile pets. Given that our brain has evolved to aid our survival and reproduction (as any biological organ), we can expect it to be geared towards churning out true beliefs. False beliefs, or cognitive mechanisms that tend to produce such beliefs, are mostly harmful and will be weeded out by natural selection.

---

Maarten Boudry is at the Department of Philosophy and Moral Sciences, Ghent University. Correspondence to: Vakgroep Wijsbegeerte en Moraalwetenschap, Universiteit Gent, Sint-Pietersnieuwstraat 49, 9000 Gent, Belgium. E-mail: [maartenboudry@gmail.com](mailto:maartenboudry@gmail.com). Michael Vlerick is at the Department of Philosophy, University of Johannesburg. Correspondence to: Department of Philosophy, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa. E-mail: [michaelvlerick@gmail.com](mailto:michaelvlerick@gmail.com)

45 According to a controversial argument developed by Alvin Plantinga (1993, 2000),  
 this assumption is false. If our cognitive faculties evolved by natural selection, he  
 argues, we have no reason to trust their deliverances. Natural selection does  
 not care about truth, but only about fitness. If evolution is blind and unguided,  
 then there is no guarantee at all that it will produce organisms with reliable, truth-  
 50 conducive cognitive faculties.

Naturalism and evolutionary theory, Plantinga contends, form a dangerous package  
 deal. Roughly, naturalism can be construed as the claim that no gods or supernatural  
 beings exist, and evolutionary naturalism as the standard scientific view of unguided  
 evolution through purely natural mechanisms. If both are true, we cannot expect our  
 55 cognitive faculties to be reliable and (the bulk of) our beliefs to be true. In the end, the  
 evolutionary naturalist must fall into an abyss of scepticism, acquiring a defeater for all  
 of his beliefs, including his very belief in evolutionary naturalism.

Earlier critics (Fales 1996, 2002) have defended the common-sense view, against  
 Plantinga, that true beliefs are generally better guides in the world than false ones,  
 60 barring contrived exceptions, and that our expensive cognitive machinery is unlikely  
 to have evolved without conferring some adaptive benefits. Stephen Law (2012) has  
 argued that there is a strong conceptual link between beliefs and behaviour, which  
 makes beliefs ‘visible’ to natural selection. Natural selection weeds out beliefs that  
 are not conducive to an organism’s survival and reproduction, among which (often  
 65 enough) false beliefs.

Plantinga’s latest formulation of his evolutionary argument against naturalism  
 (EAAN) does not differ much from earlier versions, but suffers from three additional  
 and important flaws, which are relevant to the constructive project of evolutionary  
 epistemology: a failure to appreciate adaptive path dependency, an incoherent  
 70 conception of content ascription, and a conflation of common sense and scientific  
 beliefs. More fundamentally, Plantinga’s rigid formalist approach to the issue of  
 cognitive reliability fails to deal with relevant empirical findings.

## 75 **2. Stark Choices**

Before documenting the three flaws in Plantinga’s argument, we should say a few  
 words about the way he frames the issue of cognitive reliability. Plantinga forces us  
 to make a stark choice between two mutually exclusive propositions, to wit  $R$  (‘our  
 cognitive faculties are reliable’) or its negation, not- $R$  (‘our cognitive faculties are  
 80 not reliable’). Having framed the discussion in those dichotomous terms, he then  
 proceeds to determine the value of  $P(R | N \& E)$ , where  $N$  is naturalism and  $E$  is  
 evolutionary theory. Serious evolutionary epistemology, however, does not frame  
 the reliability of human cognition in such stark, black-or-white terms. If only it  
 were that simple. Relevant empirical complications include the structure of the  
 85 ancestral environment, various design trade-offs involving production costs and  
 error rates, the specific adaptive problem at hand, the specific selection pressures  
 and developmental constraints at work, etc. For example, evolution may endow us  
 with ‘fast and frugal’ heuristics (Gigerenzer 2008) that sacrifice some accuracy for

90 efficiency and speed; evolved cognitive mechanisms may produce occasional false belief as a result of asymmetries in the costs associated with different errors (Haselton and Buss 2000); cognitive mechanisms may be accurate or truth-conducive in adaptively familiar environments but lead us astray in new or modern environments; adaptive beliefs may be roughly accurate approximations of the world though limited in other respects. Evolution may select true belief in some domains, but not necessarily in others (De Cruz et al. 2011; Wilkins and Griffiths 2012). However, misbelief that is *directly* adaptive turns out to be hard to find (McKay and Dennett 2009). In general, the standard view of evolutionary epistemology was neatly summarized in a seminal paper by Ryan McKay and Daniel Dennett on the possibility of adaptive misbelief: 'Although survival is the only hard currency of natural selection, the exchange rate with truth is likely to be fair in most circumstances' (McKay and Dennett 2009, 509).

100 The strand of reductive formalism that Plantinga promotes is—to put it mildly—unlikely to yield a fruitful philosophical discussion, because it sweeps many empirical details and qualifications under the carpet. Many philosophical commentators, defenders of Plantinga and critics alike, have indulged in his reductionist framework, instead of paying attention to the actual empirical work that is being carried out at the moment. In James Beilby's volume devoted to the EAAN (Beilby 2002), most contributors (see Talbott, O'Connor, Bergmann, Sosa, and others) accept Plantinga's black-or-white distinction between  $R$  and not- $R$ . They even grant the low value of  $P(R | N \& E)$ , for the sake of the argument, and prefer to quarrel over various probabilities, the exact formalization of the argument, and a whole network of defeaters and defeater-deflecting defeaters (Evan Fales's chapter is a notable exception). But that is to give the game away right from the start. The only apposite response to Plantinga's formal dichotomy is the one voiced by Geoff Childers, in another critical paper on the EAAN: 'Such cartoonish skepticism arises only as a byproduct of forcing ourselves to think in artificially rigid black and white categories, while ignoring our best current naturalistic models of the biological and social evolution of cognition' (Childers 2011, 201).

### 120 3. Evolutionary Scepticism

In this article, we will not accept Plantinga's 'cartoonish' framing, but instead focus on a number of specific and important flaws underlying his EAAN, which have not been covered yet in the literature and which are relevant for the project of evolutionary epistemology. Helen De Cruz et al. (2011) have distinguished between evolutionary supportive arguments (ESA), which ground the truth-tracking abilities of our cognitive faculties in their evolutionary origins, and evolutionary debunking arguments (EDA), which try to sever or block at least one of the steps in the supportive ESA. The general framework of the ESA runs roughly like this:

- 130 (1) Organisms need to interact with the environment to survive and reproduce (premise)  
 (2) An organism's beliefs guide its interaction with the world (premise)

- (3) There is a correlation between an organism's fitness (likelihood of survival and reproduction) and its beliefs (from 1 and 2)
- 135 (4) True beliefs are overall better guides to action in the world than false beliefs (premise, see above for subtle empirical qualifications)
- (5) Beliefs have a neural substrate (premise)
- (6) Natural selection will weed out neural structures that give rise to false beliefs (from 3, 4, and 5)
- 140 (7) Natural selection is the main driving force of evolution (premise)
- (8) Evolution will produce organisms with cognitive systems that produce (overall) true belief (from 6 and 7)
- (9) Our belief in evolutionary naturalism is justified

145 One can distinguish different types of EDA depending on which of these steps in the ESA are being blocked or challenged. Some philosophers have attacked step 7 in particular (Stich 1990). Plantinga's case is more focused around step 4, although he has taken issue with other steps as well. According to the EAAN, evolution by natural selection does not care about truth; it cares only about taking adaptive courses of action. There is no reason to think that we need true belief to achieve this. Maybe

150 our beliefs are causally inert, or they are mere by-products of adaptive behaviour, or maybe false beliefs can be adaptive.

In the latest formulation of his argument, Plantinga commits a number of flaws that are relevant for the project of evolutionary epistemology. The first reasoning error discussed in this article consists of an attempt to block step 4, whereas the second targets

155 step 5. The third flaw concerns step 9. Even if the ESA is not airtight up to step 8, meaning that our cognitive systems are not *completely* reliable, we can still have confidence in our belief in evolutionary naturalism. Plantinga's conflation of different sorts of belief is described as the 'foundationalist fallacy'.

#### 160 4. Adaptive Path Dependency

Does survival really require true or accurate belief? Many organisms, Plantinga points out, seem to do perfectly fine without it. For example, most primitive organisms have

165 simple neural structures processing input from sense organs. These 'indicators' may be adaptive, if they are correlated with specific courses of action (e.g. fleeing in the presence of a predator), but they need not involve belief states: 'indication of this sort does not require *belief*. In particular, it does not require belief having to do with the state of affairs indicated' (Plantinga 2011b, 328). It is true, of course, that survival does not

170 require belief: bacteria and other lower organisms thrive well without it. Even if organisms have evolved neural architecture, their cognitive mechanisms may be relatively simple and not require full-blown belief. But Plantinga's turn of phrase reveals a subtle error: 'Fleeing predators, finding food and mates—these things require cognitive devices that in some way track crucial features of the environment, and are appropriately connected with muscles; but they do not require *true* belief, or even belief at

175 all' (Plantinga 2011b, 329; our italics).

Plantinga equivocates between two different claims: (1) biological fitness does not require beliefs (and hence *a fortiori* no true beliefs); (2) given that an organism has action-directing beliefs, those beliefs need not be true. The second claim, however, does not follow from the first, because it ignores adaptive *path dependency* (Boudry 2013). Though in general survival and reproduction need not require (full-blown) belief, once evolution has gone down the road of belief generation as a means of directing action, it better be producing beliefs—or so the evolutionary naturalist claims—that are at least approximately true in ecologically relevant situations (and this is where the empirical details come in).<sup>1</sup> The existence of creatures who thrive without the help of beliefs *at all* is irrelevant to the claim, defended by the evolutionary naturalists, that action-directing beliefs, once adopted, are adaptive if true and harmful if false.

Adaptive path dependency is rampant in the living world. For example, most organisms do not rely on powered flight to survive, so they can do without wings. Most birds are dependent on wings, however, because their ancestors have ‘invested’ in the strategy of flight. The adaptive value of wings is not diminished by the fact that most other organisms make a good living without them. Early choices limit later options.

## 5. Arbitrary Content Labelling

If an organism has invested in neural machinery for action-directing beliefs, those should better be approximately true beliefs (Fales 2002). Is the link between true belief and fitness then secured? Not according to Plantinga, who argues that we still have to bridge the gap between the neurophysiological (NP) properties of beliefs and their contents: ‘The neurology causes adaptive behavior and also causes or determines belief content: but there is no reason to suppose that the belief content thus determined is true’ (Plantinga 2011b, 327). For the evolutionary naturalist, however, the content of a belief is associated with certain NP properties.<sup>2</sup> A belief with a different content is linked with different physical properties (i.e. a different physical realization in the brain), and will thus elicit different kinds of actions, some of which will be fitness enhancing. Natural selection selects those beliefs that result in adaptive courses of action.

There need not even be a *causal* link between belief content and behaviour. As Law (2012) has shown, constraints on belief content obtain even if (semantic) epiphenomenalism or substance dualism were true. The link obtaining between belief content and behaviour is entirely conceptual, not dependent on any specific reductionist account of meaning. For example, for a dualist

if a given soul-stuff structure in combination with a strong desire for water typically results in subjects walking five miles south, then *ceteris paribus* that soul-stuff structure is quite likely to have the content that there’s water five miles south, and is rather unlikely to have the content that there’s water five miles north. (Law 2012, 46)<sup>3</sup>

As long as NP states are associated with specific belief contents, either causally or conceptually, belief content is within the reach of natural selection, and Plantinga’s EAA

fails. The precise relation between belief content and NP properties is irrelevant to Plantinga's argument.

Plantinga objects to this conceptual link between NP properties and content, using an intriguing counterfactual argument that involves a second flaw:

225 It is exceedingly difficult to see, however, how they [beliefs] can enter that chain *by virtue of their content*. A given belief, it seems, would have had the same causal impact on behavior if it had had the same NP properties, but different content. (Plantinga 2011a, 436; see also Plantinga 2011b, 335–337)

230 If there is a conceptual link between NP properties and belief content, however, the question *in virtue of which* property beliefs cause certain courses of action becomes moot, as far as the visibility for natural selection is concerned. Because of the conceptual links, evolution will select for belief contents *even if* they are causally inert. Plantinga implicitly denies this conceptual link, however, as he asks us to imagine a NP state with identical causal powers but 'quite different content' (Plantinga 2002, 214).  
235 If one accepts that beliefs supervene on NP properties, however, this will be ruled out. For the physicalist, it is simply impossible that there can be difference in content without difference in NP properties.

In a telling passage, Plantinga (2002, 337) claims that the NP properties of a concrete belief stand to its content as a ball shattering a window stands to its being a birthday present. The content is thus completely irrelevant. Does a ball break a window 'by virtue of being a birthday present', asks Plantinga rhetorically? Of course not: some balls are not birthday presents, and some birthday presents are not balls but rather, say, stuffed animals, which usually do not shatter windows. Unlike belief contents and NP properties, the property of being a birthday present is not correlated to its breaking force and is thus irrelevant.  
245

Plantinga's counterfactual argument assumes that any belief content can be associated with any NP state, as if it were an arbitrary label: 'would it not produce the same physical effect, the same effect on nerves and muscles and glands, even if it had quite different content?' (Plantinga 2002, 214). Elsewhere, he writes that 'those beliefs need not be so much as *about* the objects involved in the state of affairs causing the subvening properties. They could be anything' (Plantinga and Tooley 2008, 232).  
250

This view—let's call it *arbitrary content labelling*—is much stronger (and stranger) than semantic epiphenomenalism. Not only does it deprive belief contents of any causal powers, but also it destroys any conceptual link with NP properties. Just as any object can assume the role of a birthday present (though not necessarily a good one), any NP structure can have the content 'there is water five miles north'. To see what Plantinga's view would amount to, consider an old-fashioned calculator. When I press some keys, numbers appear on the screen. Now suppose that Plantinga would come along and say that 'the calculator would produce the same result if the keys had different meaning'. By way of demonstration, he removes the labels from keys 3 and 4 and switches them: 'see, the calculator still produces the correct results'. But it should be obvious that the inscription or label on the key is not where meaning resides. The pressure applied to the key relays a signal on an electronic  
255  
260

265 circuit, regardless of the inscriptions on the key. The key represents number 3 in virtue  
of the fact that it is connected with the device in a certain manner, such that it pro-  
duces the expected results on the screen when pressed in combination with other  
keys. The manufacturer has just conveniently arranged my keypad in such a way  
270 that the inscriptions correspond to the internal representations, in this case  
numbers and mathematical concepts. If somebody secretly swapped the key labels,  
this will soon be discovered ('this key is not "3" at all!').

Believing *X* means being prepared to take certain courses of action, when conjoined  
with a set of desires and other beliefs. As William Clifford already wrote: 'no belief is  
real unless it guides our actions, and those very actions supply a test of its truth' (Clif-  
275 ford [1877] 1999, 82). If some NP property does not dispose us, *ceteris paribus*, to the  
behaviour that we would expect from some belief state *X*, then *it cannot be that belief  
state X*. For instance, if the belief 'this snake is dangerous' does not produce flight be-  
haviour in the presence of a snake, when coupled with the desire 'I don't want to get  
280 bitten by a poisonous snake', and with other background beliefs, such as 'if I run away  
from *X*, then *X* is less likely to hurt me', then something must have gone wrong in our  
belief ascription.<sup>4</sup> Belief contents are not arbitrary labels that one can tag to NP prop-  
erties at will, as Plantinga supposes. The counterfactual claim that some NP property  
would have the same effect 'even if it had quite different content', therefore, is  
incoherent.<sup>5</sup>

285 Is that all there is to be said about Plantinga's severing of behaviour and semantics?  
No, because on other occasions, Plantinga slides from a counterfactual argument  
about *content*, which suffers from incoherence, into a superficially similar counterfac-  
tual argument centring on *truth-values* instead of belief contents. In this version, he  
accepts a form of reductive materialism, which simply identifies certain NP properties  
290 with belief contents. This ensures that evolution can track *adaptive* beliefs. Fair  
enough, argues Plantinga, but does this give us any confidence that such beliefs are  
also *true*?

Take a neurophysiological property *P* that is in fact adaptive, and is also identical  
with the property of having content *p*: so far as adaptivity goes, it doesn't in  
295 general matter whether *p* is true or false. *P* would have the same adaptive effects,  
one thinks, if *p* were false. How *P* contributes to motor output doesn't depend  
(except in special cases) upon the truth-value of *p*. (Plantinga 2002, 218)

In this version, Plantinga accepts the disputed conceptual link between content and  
NP properties: 'The neurology causes adaptive behavior and also causes or determines  
300 belief content: but there is no reason to suppose that the belief content thus deter-  
mined is true' (Plantinga 2011b, 327). It is important to see that this is a different argu-  
ment from the one above about belief *content*. Indeed, this argument is hardly  
incoherent: a given NP property may instantiate a belief that is either true or false,  
depending on the state of the world. The belief of an early hominid that 'this mush-  
305 room is not poisonous' has that particular content in virtue of his neurological state  
(within a particular environment). The NP properties of *p* will result in the same type  
of behaviour (eating that tasty mushroom), regardless of the truth-value of *p*. By  
themselves, the NP properties of *p* do not give us any information about the *truth-*

310 *value* of  $p$ . If  $p$  is false, however, because the organism failed to identify the type of mushroom, the NP properties linked to  $p$  will fail to be adaptive, *pace* Plantinga.<sup>6</sup>

The link that the evolutionary naturalist seeks to secure is not between *truth* and course of action, but between truth and *adaptiveness*. There are two connections at play: the straightforward one between belief content and course of action, and the higher-order connection between truth and adaptiveness. Plantinga muddies the waters by mixing up these two levels, alternately asking what would happen if a certain NP property had a different *content*, and what if it had different *truth-value*. The two different links need to be kept apart: (1) The NP properties of my belief ‘there is a bear in the cave’ is linked with my disposition not to enter it, *regardless* of whether that belief is true. (2) The adaptiveness of the NP properties that dispose me not to enter the cave, under those circumstances, depends on the state of the world (the bear’s presence), *regardless* of the semantic content of those NP properties. By taking either of those links in isolation, and switching back and forth between them, Plantinga seems to undermine the confidence of the evolutionary naturalist. There always seems to be a ‘missing link’. But appearances are misleading. Taken together, however, those claims provide all the necessary chains to establish a link between fitness and truth.

## 6. The Foundationalist Fallacy

330 Remember that Plantinga ends up rejecting step 8 of the ESA, by blocking one of the earlier steps:

(8\*) Evolution will NOT produce organisms with cognitive systems that produce (overall) true belief

335 From this assessment of our ‘naked’ and unassisted cognitive faculties, given evolutionary naturalism, Plantinga immediately jumps to the conclusion that scientific beliefs—notably, belief in naturalistic evolution—must also be unreliable. In its basic form, therefore, the punch line of the EAAN goes as follows:

- 340 (1) Given evolutionary naturalism, our belief producing cognitive faculties are unreliable  
 (2) Evolutionary naturalism is itself a product of our cognitive faculties  
 (3) Conclusion: evolutionary naturalism is self-defeating

Indeed, Plantinga argues,

345 science, clearly, is an extension of our ordinary ways of learning about the world. As such, it obviously involves the faculties and processes by which we ordinarily do achieve knowledge. . . . For science to be successful, therefore, there must be a match between our cognitive faculties and the world. (Plantinga 2011b, 269–270).

350 Once again, Plantinga sets up a stark dichotomy: either all our beliefs are reliable or none of them is, including scientific beliefs. More generally, he conflates beliefs that are the direct product of our naked, unassisted mind, and beliefs that arise out of higher-order cognitive faculties (systematic theorizing and evidence-gathering).

Underlying this argument is the assumption, which we call the ‘foundationalist fallacy’, that if the foundations of our cognitive endeavours are not *completely* secure, then the whole edifice built on top of it must collapse. If the ground is even a *little* shaky, no amount of scaffolding will be sufficient.

As John Clendinnen (1989) has argued, however, epistemic confidence can be gradually raised as evidence accumulates. Consider the triplet of propositions ‘If *P* then *Q*’, ‘If *Q* then *R*’, ‘If *R* then *P*’. In a deductive model of justification, it is often incorrectly assumed that each step is either completely justified or not justified at all. However, it is possible to have an interdependence of justifications (*P*, *Q*, *R*) in which we start out by an initially very weak and provisional version of *P*, whereby our confidence is gradually raised through *Q* and *R*. In this regard, by starting from the deliverances of our bare and unassisted mind, including folk logic and basic perceptual capacities, we may gradually increase the reliability of our beliefs, by accumulating evidence and thus bootstrapping or ‘scaffolding’ (Sterelny 2010) our way to more justified beliefs. Therefore, even if we grant that step 8 of the ESA is not valid across the board (see the empirical complications above), the EAAN’s conclusion does not follow. Indeed, in his rigidly foundationalist approach, Plantinga leaves no room for such gradual bootstrapping operations.

Let us have a closer look at the belief in evolutionary naturalism. The theory of evolution by natural selection is by no means the product of our bare, unassisted mind. It is in every respect the result of a long process of critical thinking and of ‘scaffolding’ our naked mind, as Kim Sterelny (2010, 480) puts it. It is supported by a massive body of knowledge accumulated over many generations (about species, fossils, genes, etc.), aided by formal scientific systems of analysis (e.g. mathematical models in population genetics), and technological innovations (e.g. microscopes in genetic research, spectrometers for the carbon dating of fossils, etc.). Quite a different kind of belief than the one held by Plantinga’s beloved prehistoric hominid Paul about the nature of the hungry tiger in front of him (Plantinga 1993, 225–226). Indeed, as Sterelny (2010, 480) points out, scientific theories about the world are primarily the product of our ability to extend our mind’s capacity.

First, the impressive body of scientific knowledge is not the product of a single mind, but the result of cooperation, both horizontally (groups of cooperating scientists rather than isolated individuals) and vertically (cumulative progress over many generations). Michael Tomasello, in this regard, points to our ability to read the minds of conspecifics as the key cognitive adaptation accounting for human culture (and by extension science). This, according to Tomasello, allows human beings ‘to pool their cognitive resources both contemporaneously and over time in ways that are unique in the animal kingdom’ (Tomasello 2001, 135). In a similar vein, Robert Boyd, Peter J. Richerson, and Joseph Henrich argue that, rather than to innate intelligence, we owe our success to ‘our uniquely developed ability to learn from others’ (Boyd, Richerson, and Henrich 2011, 1).

Second, our best scientific theories are supported by a series of cognitive artefacts (e.g. mathematics, providing us with a formal system to represent and process quantitative information) and various notation devices, enabling us to store huge amounts

of information and work out complex mathematical equations (De Cruz 2008). The leverage provided by these scaffolds can hardly be overestimated. They radically transform our ability to retain, transmit and correct knowledge about the world. Finally, and equally important, we have a myriad of technological instruments at our disposal, enabling us to detect data we could never gather by means of our naked senses (e.g. telescopes, microscopes, barometers, and the like), and process information in ways we could never do with our naked mind (computers).

Even on the level of individual belief formation, apart from the social organization of science, we should distinguish between different types of beliefs. Daniel Kahneman (2011) and other researchers of dual process theory have shown that, while the fast, intuitive and almost automatic belief outputs of what he dubbed ‘system 1’ can occasionally—and in some cases even predictably—lead us astray, we do possess the cognitive ability to override these beliefs (Vlerick 2012). ‘System 2’ does indeed provide us with a slower, voluntary and more strenuous cognitive activity, which avoids many of the pitfalls of system 1. How else could we identify sloppy and biased reasoning arising from system 1?

In this perspective, while natural selection may—as pointed out (cf. section 2)—provide us with the occasional false intuitive belief (e.g. as the by-product of fast and frugal heuristics or, in Kahneman’s terms, the output of system 1), this obviously does not apply to beliefs that are supported by elaborate cognitive scaffolds and interconnected strands of evidence, such as the belief in evolution by natural selection. This prompts Michael Levin (1997, 98) to conclude that ‘there is no evidence that evolution has fated us to theoretical error, *particularly* about evolution itself’ (our italics). Indeed, science can be seen as a system of safeguards against our intuitive biases and misconceptions, or what Francis Bacon called the ‘idols of the mind’ in his *Novum Organum*. As Levin (1997, 97) points out, even if our cognitive faculties mislead us, ‘we can find this out and correct for it by comparing the predictions of our theories to observation’.

As it turns out, the belief in evolution by natural selection is a textbook case in overcoming our intuitive grasp of the world. Notably, Darwin’s theory violates our intuitive species essentialism (a useful heuristic with which evolution endowed us to make ‘quick and dirty’ inferences about groups of living organisms), which predisposes us to view species as eternal and endowed with an immutable essence. Population thinking, which came to replace essentialism in biology, views species as loose aggregates of individuals showing a certain distribution of characteristics rather than an immutable set of properties. Moreover, the creative power of blind selection also conflicts with our teleological intuitions (in the sense of goal-directedness) about functional complexity, and our intuitive design stance (Kelemen and Rosset 2009; Reiss 2009).

Ironically enough, creationism in particular and theism in general are typical cases of belief systems that seem to entrench our biased, intuitive ways of understanding the world (McCauley 2011; Blancke et al. 2012). In this regard, such beliefs would be the first to succumb to a sceptical attitude towards the outputs of adaptive cognitive modules. Thus, if anything, Plantinga’s EAAN backfires against theism. As Childers (2011, 199) notes: ‘the atheist has a license to wonder whether theistic belief is one

of those false, yet adaptive cognitive dispositions which, as Plantinga notes, would tend to thrive under natural selection’.

## 7. Conclusion

Plantinga’s latest formulation of the EAAN contains three major flaws. The first is based on a failure to appreciate adaptive path dependency; the second derives from his confused account of belief content and NP (neurophysiological) properties, and the third from an unwarranted conflation of common sense and scientific beliefs. Uncovering these reasoning errors, we have aimed to show how Plantinga’s reductive formalism, by forcing a stark choice on the evolutionary epistemologist, leads to scholastic discussions that are out of touch with the relevant empirical literature. It is the philosophical equivalent of performing brain surgery with an axe. Natural selection is not the Cartesian evil genius Plantinga takes it to be, nor has it condemned us to be slaves to our intuitive cognitive modules. Evolution *does* care about truth. Besides providing us with overall reliable beliefs, it has endowed us with the ability to ‘scaffold’ our way up to systems of beliefs that are even more reliable, accurate, and empirically successful. But is theism among them?

## Acknowledgements

We would like to thank Stephen Law, Eric Schliesser, Herman Philipse, and Alex Broadbent for commenting on an earlier version of this paper. Thanks also to two anonymous reviewers, and to James McAllister, for their valuable comments.

## Notes

- [1] For the time being, we take for granted that belief formation directs behaviour in the first place (premise 2). As Fales (1996) has argued, the cognitive machinery for belief formation is energy-consuming for the organism, and is unlikely to have evolved without making a difference in behaviour. Beliefs are not on a permanent holiday.
- [2] Indeed, the natural properties realizing content need not be (solely) neurophysiological. If one takes an externalist point of view, for example, content would be determined by the physical state of the brain-plus-environment. For the sake of brevity, however, we will keep on referring to NP properties as instantiating meaning.
- [3] There are always contrived ways of avoiding this conclusion, for example, by imagining an agent with an internal compass that has the wind directions exactly in reverse.
- [4] This assumes a (moderate) form of semantic holism: beliefs ‘come to life’ in relation to other mental states, and cannot be interpreted in splendid isolation.
- [5] Plantinga (2011b, 338) has a reply of sorts to this objection, but it forces him to call the whole semantics of counterfactuals into question. His alternative proposal accepts reasoning about counterpossibles (‘if 2 had been greater than 3’). This, for instance, would permit him to say things such as ‘If 13 were divisible by a natural number apart from one and itself, it would not be a prime number’. This is already quite strange. So is it ‘in virtue of’ not having any positive divisors, other than 1 and itself, that 13 is a prime number? Or the other way around? But this is to talk nonsense. In any case, Plantinga writes that he will not open ‘that particular (and large) can of worms’. So neither shall we.

- 485 [6] For a critique of Plantinga's far-fetched scenarios where corrupt beliefs happen to result in felicitous behaviour, see Fales (2002).

## References

- 490 Beilby, J. K. 2002. *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*. Ithaca, NY: Cornell University Press.
- Blancke, S., J. De Smedt, H. De Cruz, M. Boudry, and J. Braeckman. 2012. "The Implications of the Cognitive Sciences for the Relation Between Religion and Science Education: The Case of Evolutionary Theory." *Science and Education* 21: 1167–1184.
- 495 Boudry, M. 2013. "Alvin Plantinga: Where the Conflict Really Lies. Science, Religion and Naturalism." *Science and Education* 22: 1219–1227.
- Boyd, R., P. J. Richerson, and J. Henrich. 2011. "The Cultural Niche: Why Social Learning Is Essential for Human Adaptation." *Proceedings of the National Academy of Sciences* 108: 10918–10925.
- Childers, G. 2011. "What's Wrong with the Evolutionary Argument Against Naturalism?" *International Journal for Philosophy of Religion* 69: 193–204.
- 500 Clendinnen, F. J. 1989. "Evolutionary Explanation and the Justification of Belief." In *Issues in Evolutionary Epistemology*, edited by K. Hahlweg and C. A. Hooker, 458–474. Albany: State University of New York Press.
- Clifford, W. K. [1877] 1999. *The Ethics of Belief and Other Essays*, edited by T. Madigan. Amherst, MA: Prometheus Books.
- De Cruz, H. 2008. "An Extended Mind Perspective on Natural Number Representation." *Philosophical Psychology* 21: 475–490.
- 505 De Cruz, H., M. Boudry, J. De Smedt, and S. Blancke. 2011. "Evolutionary Approaches to Epistemic Justification." *Dialectica* 65: 517–535.
- Fales, E. 1996. "Plantinga's Case Against Naturalistic Epistemology." *Philosophy of Science* 63: 432–451.
- Fales, E. 2002. "Darwin's Doubt, Calvin's Calvary." In *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument Against Naturalism*, edited by J. K. Beilby, 43–58. Ithaca, NY: Cornell University Press.
- 510 Gigerenzer, G. 2008. *Rationality for Mortals: How People Cope with Uncertainty*. Oxford: Oxford University Press.
- Haselton, M. G., and D. M. Buss. 2000. "Error Management Theory: A New Perspective on Biases in Cross-sex Mind Reading." *Journal of Personality and Social Psychology* 78: 81–91.
- 515 Kahneman, D. 2011. *Thinking, Fast and Slow*. London: Penguin Books.
- Kelemen, D., and E. Rosset. 2009. "The Human Function Compunction: Teleological Explanation in Adults." *Cognition* 111: 138–143.
- Law, S. 2012. "Naturalism, Evolution and True Belief." *Analysis* 72: 41–48.
- Levin, M. 1997. "Plantinga on Functions and the Theory of Evolution." *Australasian Journal of Philosophy* 75: 83–98.
- 520 McCauley, R. N. 2011. *Why Religion Is Natural and Science Is Not*. New York: Oxford University Press.
- McKay, R. T., and D. C. Dennett. 2009. "The Evolution of Misbelief." *Behavioral and Brain Sciences* 32: 493–561.
- Plantinga, A. 1993. *Warrant and Proper Function*. New York: Oxford University Press.
- Plantinga, A. 2000. *Warranted Christian Belief*. New York: Oxford University Press.
- 525 Plantinga, A. 2002. "Reply to Beilby's Cohorts." In *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument Against Naturalism*, edited by J. K. Beilby, 204–275. Ithaca, NY: Cornell University Press.
- Plantinga, A. 2011a. "Content and Natural Selection." *Philosophy and Phenomenological Research* 83: 435–458.

- 530 Plantinga, A. 2011b. *Where the Conflict Really Lies: Science, Religion, and Naturalism*. New York: Oxford University Press.
- Plantinga, A., and M. Tooley. 2008. *Knowledge of God*. Oxford: Blackwell.
- Reiss, M. J. 2009. "The Relationship Between Evolutionary Biology and Religion." *Evolution* 63: 1934–1941.
- Sterelny, K. 2010. "Minds: Extended or Scaffolded?" *Phenomenology and the Cognitive Sciences* 9: 465–481.
- 535 Stich, S. P. 1990. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press.
- Tomasello, M. 2001. "Cultural Transmission: A View from Chimpanzees and Human Infants." *Journal of Cross-cultural Psychology* 32: 135–146.
- Vlerick, M. 2012. "How Can Human Beings Transgress Their Biologically Based Views?" *South African Journal of Philosophy* 31: 707–735.
- 540 Wilkins, J. S., and P. E. Griffiths. 2012. "Evolutionary Debunking Arguments in Three Domains." In *A New Science of Religion*, edited by G. Dawes and J. Maclaurin, 133–146. London: Routledge.

545

550

555

560

565

570