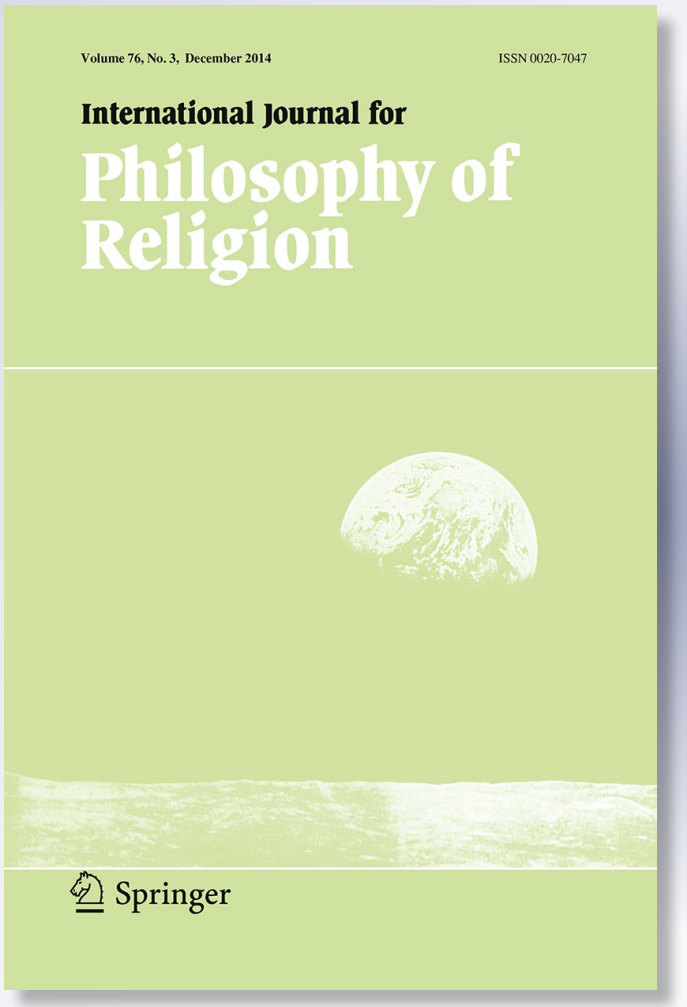
Evolutionary debunking arguments against theism, reconsidered

Jonathan Jong & Aku Visala

International Journal for Philosophy of Religion

ISSN 0020-7047
Volume 76
Number 3

Int J Philos Relig (2014) 76:243-258
DOI 10.1007/s11153-014-9461-6



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Author's personal copy

Int J Philos Relig (2014) 76:243–258
DOI 10.1007/s11153-014-9461-6

ARTICLE

Evolutionary debunking arguments against theism, reconsidered

Jonathan Jong · Aku Visala

Received: 4 February 2014 / Accepted: 25 April 2014 / Published online: 11 May 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Evolutionary debunking arguments (EDAs) against religious beliefs move from the claim that religious beliefs are caused by off-track processes to the conclusion that said religious beliefs are unjustified and/or false. *Prima facie*, EDAs commit the genetic fallacy, unduly conflating the context of discovery and the context of justification. In this paper, we first consider whether EDAs necessarily commit the genetic fallacy, and if not, whether modified EDAs (e.g., those that posit falsehood-tracking or perniciously deceptive belief-forming mechanisms) provide successful arguments against theism. Then, we critically evaluate more recent attempts to argue that a more promiscuous evolutionary scepticism renders religious belief unjustified because, unlike commonsense and scientific beliefs, religious beliefs have no way out of such scepticism.

Keywords Evolutionary debunking · Religion · Cognitive science of religion · Epistemology

Evolutionarily debunking arguments

There has recently been some excitement regarding the application of so-called "evolutionary debunking arguments" (EDAs; Kahane 2011) as arguments against theism (cf. Cruz et al. 2011; Griffiths and Wilkins 2013; Leech and Visala 2011, 2012). The

J. Jong (✉)
Institute of Cognitive and Evolutionary Anthropology, University of Oxford,
64 Banbury Road, Oxford OX2 6PN, United Kingdom
e-mail: jonathan.jong@anthro.ox.ac.uk

A. Visala
Department of Anthropology, University of Notre Dame,
611 Flanner Hall, Notre Dame, IN 46556-5611, USA

Author's personal copy

suggestion that evolutionary explanations of certain beliefs might undermine those beliefs is hardly new, of course; Darwin (1881) himself famously worried about the ability of human cognition to arrive at metaphysical truths for evolutionary reasons:

> But then with me the horrid doubt always arises whether the convictions of man's mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy. Would any one trust in the convictions of a monkey's mind, if there are any convictions in such a mind?

Following Darwin's doubt, there has been a substantial and varied body of philosophical discussion on evolutionary scepticism in epistemology (e.g., Plantinga 1993; see Beilby 2002 for extended discussion) and ethics (e.g., Brosnan 2011; Joyce 2006); these more recent applications of EDAs in philosophy of religion are just extensions of this research programme, following predictably in the footsteps of the slightly less recent enthusiasm for evolutionary explanations of religion (e.g., Atran 2002; Barrett 2004; Bering 2011; Boyer 2001; Dennett 2006; Tremlin 2006; Wilson 2002).

This paper aims to critically evaluate the effort to construct EDAs against theism and other religious beliefs, first by scrutinizing EDAs (and other *genealogical* debunking arguments) more generally and then by turning to more specific issues regarding anti-theistic EDAs. As a starting point, Kahane (2011) conveniently supplies a general schema for debunking arguments, as follows:

> *Causal premise*. S's belief that p is explained by X
> *Epistemic premise*. X is an off-track process
>
> [where an "off-track process" is just one that does not track the truth of p or p-type beliefs[1]]
>
> Therefore
> S's belief that p is unjustified

There are a few features of this general schema to just briefly note. First, the argument concludes with a claim about whether or not p is *justified*, not whether or not p is *true*. There have been attempts to construct EDAs against moral beliefs that conclude that our moral beliefs are probably false (e.g., Ruse 2006, 2009; Street 2006), but this just seems odd, especially under a correspondence theory of truth and a realist semantics of moral propositions. If moral propositions are true or false regardless of whether *anyone* believes them to be true or false, surely *how* they come to believe them to be true or false is equally irrelevant to their truth values. Nevertheless, we shall look at EDAs that conclude with the (probable) falsity of p later, even though

[1] This raises the question of what truth-tracking consists in. Kahane (2011) does not specify, instead assuming "an intuitive understanding of the epistemic premise (such an understanding doesn't require a positive account of truth-tracking processes)". Presumably, however, he has something like Nozick's (1981) account in mind, in which:

(1) p is true
(2) S believes p
(3) if p were not true then S would not believe that p, and
(4) if p were true then S would believe that p (and would not believe that not-p).

Author's personal copy

this departs from Kahane's (2011) schema. The second note, to which we shall return in the latter half of this paper, is that EDAs are always going to face the challenge of establishing the epistemic premise without begging the question or trivializing the argument. EDAs have to specify what counts as a p or p-type belief, such that X is off-track with respect to these beliefs; furthermore, the relevant domain must be spelt out in a non-question begging way. It is, to pre-empt ourselves for a moment, unclear that this can be done in the case of religious beliefs, if what we mean by that is beliefs about certain kinds of agents (viz., supernatural agents). With these two promissory notes, we shall move on to a more obvious complaint against genealogical debunking arguments (evolutionary or otherwise), hitherto surprisingly under-discussed; that is, that it commits the genetic fallacy.

EDAs and the genetic fallacy

The genetic fallacy is just "the error of treating items in the context of discovery as if they belonged to the context of justification" (Salmon 1984). It is somewhat unfashionable these days to draw this distinction, but it is no less valid for being unpopular. There are many more or less specific ways to characterize the distinction—as pertaining to the subjective and objective (Reichenbach 1938), the descriptive and normative (Hoyningen-Huene 2006), psychology and epistemology (Reichenbach 1938), history of science and philosophy of science (Salmon 1970), and so forth—but, for our purposes, a more general definition will do. The context of discovery pertains to how one comes to believe something, the source or origin of the belief; the context of justification pertains to how one comes to prove, defend, or otherwise justify the belief, the arguments and evidence for it. Applied to religion, it is Hume's (1757/2008) distinction between the "origin [of religion] in human nature" and its "foundation in reason". The genetic fallacy is then the fallacy of considering facts about how a belief is formed as relevant, *ipso facto*, to whether one ought to hold that belief. The underlying recognition is that doxastic conditions do not, *by themselves*, have epistemic implications. To be sure, this is not to say that doxastic conditions are wholly epistemically irrelevant. Indeed, as process reliabilists argue, there is at least a strong inductive case for believing to be true the deliverances of reliable belief-forming mechanisms (e.g., Goldman 1986, 2008; see also BonJour 1980). Still, even in this case, the reliability of the belief-forming mechanism fails to guarantee the truth of the belief formed. Much less clear is whether or not the converse case can be made, for the move from an "off-track" belief-forming process to the falsity or unjustifiability of said belief. Debunking arguments, as schematized earlier, look very much like just such an attempt to draw negative epistemic implications simply from doxastic conditions: they conclude that "S's belief that p is unjustified" from the fact that S's belief p is explained by a doxastic process, X, that is not truth-tropic. So, what are we to make of this? Is genetic reasoning not fallacious after all? Or do debunking arguments exploit some kind of loophole or exception to an otherwise acceptable rule, and therefore only work under some restrictive conditions?

The first way we shall consider to exculpate EDAs from the charge of being fallacious is to show that the genetic fallacy is not a fallacy after all by showing that

Author's personal copy

the context distinction cannot be made. At first blush, this seems an unlikely strategy. A well-known and potentially apocryphal (or at least hyperbolical) tale from the history of science demonstrates that the context distinction can indeed be made for the purposes of the genetic fallacy: August Kekulé recounts discovering the chemical structure of benzene in the dimly-lit study of his bachelor quarters in Ghent; daydreaming as his writing was stagnant, he saw a snake seizing its own tail and behold! he "recognizes truth without knowing the evidence for it" (Benfey 1958 p. 22). This, then, was the context of discovery. Upon receiving this revelation, Kekulé worked out the implications of his new theory, and collected empirical evidence for it. This is the context of justification. Now, it seems uncontroversial to say that the fact that Kekulé's original idea came from a dream—which, we will assume, is an unreliable way of discovering chemical structures—is irrelevant to the question of whether we ought to believe that benzene has a ring structure or not. This belief is hardly irrational despite its questionable provenance; it is not irrational precisely because we can provide evidence and construct persuasive arguments in support of the belief. To think otherwise would be fallacious. It would commit the genetic fallacy (cf. Jong 2012).

The Kekulé example shows that there are conditions under which the inference from the psychological origins of a belief to the epistemic status of that belief is invalid, namely conditions under which S has good *reasons* for believing p, or under which reasons feature among the causes of S's believing p. Note that it does so regardless of whether this story is true, or just a thought experiment. Furthermore, the example does not rely on a *temporal* distinction between the contexts of discovery and justification respectively, but a conceptual one. Nor does it rely on a complete decoupling of the contexts: the general principle—that we ought not infer the epistemic status of a belief from the doxastic conditions of its genesis—stands up in the face of counterexamples to the distinction. Of course items in the context of discovery sometimes also show up as items in the context of justification; for example, we sometimes arrive at beliefs precisely because we stumble upon good reasons for them. Even so, just because the context of justification is *sometimes* identical with the context of discovery does not allow us to conflate the contexts in novel cases (see also Siegel 1980). On any reasonable understanding of justification—including externalist theories—if we have good reasons for holding a belief, then it is justified;[2] the questionable psychological origins of the belief are neither here nor there.

Perhaps EDAs are not concerned with situations like Kekulé's. Perhaps they are only concerned with situations in which, for example, Kekulé cannot provide evidence and construct persuasive arguments for the theory that benzene has a ring structure. That is, perhaps EDAs are only concerned with cases in which S has *no reasons* to believe p. One way of construing this would be to say that good reasons do not feature at all as causes of S's belief that p; S's belief that p not only *originates* in an off-track process, but it is also wholly sustained by an off-track process (or a collection of off-track processes). If so, EDAs would have very limited application, to rather unreflective epistemic agents, but still, it is interesting to consider how EDAs are meant to work even in these restrictive cases. The question remains: in what sense would a complete

[2] Indeed, for epistemic externalists, reasons could count as causes; *good* reasons would thus be epistemically respectable causes

Author's personal copy

explanation of S's belief that p in terms of an off-track process X *debunk* S's belief that p? One possibility is that such an explanation would show that p is false or probably false. Of course, a belief can still be false and yet justified; however, as Kahane (2011) observes, on any plausible view of justification, S would be unjustified in holding p once the fact that p is (probably) false is brought to her attention. However, as mentioned earlier, this argument seems unlikely to be successful on a correspondence theory of truth and a realist semantics of propositions: propositions are true or false independent on whether and why epistemic agents believe them to be true or false. A Bayesian analysis of this scenario might elucidate the point further (cf. Brosnan 2011).

If to debunk S's belief that p is to show that p is more likely false than true, the relevant Bayesian calculation would have to show that, all things considered, the posterior probability that not-p exceeds the posterior probability of p. So, let's consider the posterior probabilities of p and not-p respectively. The posterior probability of p, given S's believing p is as follows:

$$\Pr(\text{p} \mid \text{S believes p}) = \frac{\Pr(\text{S believes p} \mid \text{p}) \times \Pr(\text{p})}{\Pr(\text{S believes p})}$$

Note that the prior probability of *p* is construed broadly here, to include all evidential considerations other than the fact that S believes p. With that in mind, the posterior probability of not-*p* given S's believing p is likewise:

$$\Pr(\text{not p} \mid \text{S believes p}) = \frac{\Pr(\text{S believes p} \mid \text{not p}) \times \Pr(\text{not p})}{\Pr(\text{S believes p})}$$

As previously stated, to debunk p, Pr(not p | S believes p) must be greater than Pr(p | S believes p). That is to say, the ratio Pr(not p | S believes p) : Pr(p | S believes p) should be greater than 1. This ratio may be calculated as follows:

$$\underset{\text{Posteriors}}{\frac{\Pr(\text{not p} \mid \text{S believes p})}{\Pr(\text{p} \mid \text{S believes p})}} = \underset{\text{Likelihoods}}{\frac{\Pr(\text{S believes p} \mid \text{not p})}{\Pr(\text{S believes p} \mid \text{p})}} \times \underset{\text{Priors}}{\frac{\Pr(\text{not p})}{\Pr(\text{p})}}$$

Now, to say that X is not truth-tracking is just to say that S is as likely to believe p, regardless of whether p is true or false. That is to say, the likelihood ratio equals 1:

$$\frac{\Pr(\text{not p} \mid \text{S believes p})}{\Pr(\text{p} \mid \text{S believes p})} = 1 \times \frac{\Pr(\text{not p})}{\Pr(\text{p})}$$

Notice that this entails is that the likelihood ratio is toothless; it is not doing any work in determining the ratio of the posterior probabilities of p and not-p given that S believes p. If X is an off-track process, the likelihood ratio above has a value equal to 1, and therefore the posterior probabilities are determined wholly by the priors, by whatever independent evidence can be adduced for p and not-p respectively. In the case of evidential ambiguity, for example in which there is neither any evidence for p or not-p, the ratio of the prior probabilities also equals 1:

Author's personal copy

$$\frac{\Pr(\text{not p} \mid \text{S believes p})}{\Pr(\text{p} \mid \text{S believes p})} = 1 \times 1 = 1$$

It follows, therefore, that the ratio of the posterior probabilities also equals 1: S should be agnostic about p. The Bayesian analysis seems to undermine S's belief that p—S should no longer believe p, and instead become agnostic about p—but this conclusion is driven entirely by the values of the priors. S is indeed unjustified in her belief that p, but it is not just (or, indeed, even largely) because X is an off-track process.

Matters might be different, however, if X is a *falsehood*-tracking process; that is, rather than just generating beliefs that are orthogonal to reality, let us consider the case in which X reliably generates false beliefs. In that case, it seems like the doxastic conditions *do* have epistemic implications. Kahane's (2011) scheme may be revised as follows:

> *Causal premise*. S's belief that p is explained by X
> *Epistemic premise*. X is a falsehood-tracking process.
>
> [where a "falsehood-tracking process" is just one that always generates false beliefs]
>
> Therefore
> S's belief that p is false.

This argument seems much more promising. Indeed, upon learning that her belief that p is explained by a falsehood-tracking process, S should believe not-p (cf. Klement 2002). The problem with this new epistemic premise, however, is that there does not seem to be any serious candidates for such a falsehood-generating cognitive process. But perhaps we can settle for slightly less. Perhaps X is not *infallibly* falsehood-tracking, but is nonetheless *perniciously deceptive*, such that X very rarely causes S to believe p when p is true, and but very often causes S to believe p when p is false. Let's go back to our Bayesian analysis. Now, to say that X is perniciously deceptive is to assign a very high value to Pr (S believes p | not p) and a very low value to Pr(S believes p | p), say .9 and .0001 respectively:

$$\frac{\Pr(\text{not p} \mid \text{S believes p})}{\Pr(\text{p} \mid \text{S believes p})} = \frac{.9}{.0001} \times \frac{\Pr(\text{not p})}{\Pr(\text{p})} = 9000 \times \frac{\Pr(\text{not p})}{\Pr(\text{p})}$$

Clearly then, if X is a perniciously deceptive cognitive process, then the fact that it is a perniciously deceptive cognitive process bolsters the relative case for not-p over p. Similarly, in the absence of any other evidence—where the ratio of the priors equals 1—p is much more likely to be false than true. Indeed, in our example, p is 9,000 times more likely to be false than true:

$$\frac{\Pr(\text{not p} \mid \text{S believes p})}{\Pr(\text{p} \mid \text{S believes p})} = 9000 \times 1 = 9000$$

The question then, is whether there are any such perniciously deceptive cognitive processes. As with falsehood-tracking processes before, it does not seem that there

Author's personal copy

are. Perhaps in the future, scientists will posit cognitive mechanisms that generate religious beliefs *especially if they are false* (and rarely if they're true), but until then, we need not worry about EDAs that posit falsehood-tracking or perniciously deceptive cognitive processes.

Things are not looking good for EDAs, or at least those that posit merely off-track cognitive processes, and not falsehood-tracking or perniciously deceptive ones. They seem toothless in the face of independent evidence and arguments for p and, even in restricted cases where S's evidence for or against p is ambivalent, they do no work to render unjustifiable S's belief that p. All else being equal, in the absence of evidence either way, S should remain agnostic about p; the fact that S's belief that p is exhaustively explained by an off-track process X neither adds nor detracts to this conclusion. Yet, various philosophers have argued that EDAs can render certain beliefs unjustifiable, under certain conditions; they can hardly be unaware of these charges against it. Kahane (2011) certainly isn't. Using Marx's claims about alienation as his example, he writes:

> It seems irrelevant whether these [Marx's] claims were causally shaped by ruthless ambition, a skin condition, or an unresolved Oedipus complex. To think otherwise, it would seem, would be to commit the genetic fallacy, to confuse causes and reasons.
>
> As a general principle, it is true that when we consider a proposition someone has put forward, we should focus on the balance of reasons in its favour, not on our adversary's biography. But this point is compatible with the narrower and unremarkable claim that, when certain conditions are met, the causal origins of a belief can reduce or even remove its justification.

But what might these conditions be? As we have seen, if S fails to provide arguments or evidence for p, her belief that p is unjustified anyway (unless her belief that p counts as a properly basic belief, if there are any such beliefs); the questionable provenance of S's belief that p isn't doing the work here. If, however, S *can* provide arguments or evidence for her belief that p, then it is justified, regardless of the psychological origins of the belief. This is true, even if her quest to construct arguments or procure evidence for p is itself caused by a convergence of epistemically unsavoury factors. Kahane (2011) also anticipates *this* objection to EDAs:

> It might be objected that even if the reasons Marx gave for his beliefs were themselves shaped by influences that are off track, there might still happen to be good reasons, even if they were not truly the (explanatory epistemic) reasons for his beliefs. We ought to engage these reasons directly. It is only if we can independently show them to be plainly bad reasons that the subsidiary task of explaining how anyone would come to endorse them might be of interest.

Just so. But Kahane (2011) goes on to say that this objection falters for various reasons. First, it "fails to distinguish the question whether someone's belief *is* justified from the question whether it *could* be justified". In other words, off-track processes exhaustively explain S's believing p: that is, S *could* have good reasons for her belief that p, but *in fact* does not; no good reasons for p feature as causes in S's believing

Author's personal copy

that p. But as we have already seen, in this case, the fact that X is off-track isn't doing the debunking work here; the same effect is achieved by simply asking S to present evidence and arguments for p, and upon her failure to do so, simply pointing out that her belief that p is thereby unjustified. Again, EDAs seem to have very limited application (especially among philosophers, presumably the main audiences of articles about EDAs, whose business it is to provide evidence and arguments for their beliefs), and do not seem to be doing any work beyond pointing out that S has never had good reasons for believing p in the first place.

Kahane's (2011) second strategy is to suggest that "at some point reasons run out"; in the case of *evaluative* beliefs, for example, Kahane (2011) claims that "ultimately we can't help but appeal to intuitions… Debunking explanations of such intuitions can leave a belief lacking both actual and alternative support". Perhaps this is true for non-evaluative beliefs too; for example, it may be true for beliefs about the existence or non-existence of God. That is, perhaps even our best arguments for (or against) theism are undergirded by intuitions, which are themselves off-track. If so, according to Kahane (2011), the fact (if it is a fact) that these intuitions are off-track is relevant in evaluating the beliefs they undergird, viz., theism and atheism. This seems to be a sensible claim, but it is unclear that it adds anything to the general philosophical task of evaluating the truth of the intuitions that often appear in the premises of our arguments. Besides, the extent to which some of these intuitions are truth-tropic—for example, the principle of sufficient reason or the principle of parsimony—is notoriously difficult to evaluate, and even this evaluation rests on intuitions that may themselves be truth-tropic or off-track. As we shall come back to later, evolutionary scepticism about basic intuitions seems to undermine, not just moral and religious beliefs, but also *scientific* and *philosophical* beliefs.

Finally, Kahane (2011) ends with a rhetorical question, "But if we conclude that the intuition that supports the belief has no epistemic force, why on earth should we look for an alternative justification?" I would have thought the answer to this question is that we should look for alternative justification because, well, that's what philosophers do: our vocation is characterized by the search for good arguments (and the criticism of bad ones).

So, that was Kahane's (2011) response to the charge of committing the genetic fallacy. Richard Joyce, who applies EDAs to moral philosophy in his *Evolution of Morality* (2006) seems to concur with Kahane's (2011) first strategy; he presents a thought experiment that is meant to refute the genetic fallacy:

> Suppose that there were a pill that makes you believe that Napoleon won Waterloo, and another one that makes you believe that he lost. Suppose also that there were an antidote that can be taken for either pill. Now imagine that you are proceeding through life happily believing that Napoleon lost Waterloo (as, indeed, you are), and then you discover that at some point in your past someone slipped you a "Napoleon lost Waterloo" belief pill. ... Should this undermine your faith in your belief that Napoleon lost Waterloo? Of course it should. ... [It] is sufficient to place your belief on the dubious list. ... Knowledge that your belief is the product of a belief pill renders the belief unjustified (or perhaps shows that it was never justified in the first place, depending on one's epistemological tastes),

Author's personal copy

> demanding that unless you can find some concrete evidence either in favor or against your belief you should cease to believe this thing – that is, you should take the antidote" (pp. 179–180).

It must be admitted that this seems to be a compelling story. If, one day, someone revealed indubitably to us that we believed that Napoleon lost Waterloo because we had taken the relevant pills, we would likely lose our faith in that belief. But should we? If so, why? Again, how is this argument meant to work? Does the knowledge that we had taken "Napoleon lost Waterloo" pills refute whatever other evidence we possessed that, in fact, Napoleon lost Waterloo? As with the Kekulé case, it does not seem that we should be bothered by this new information if we have good independent reasons for believing that Napoleon lost Waterloo. Indeed, Joyce (2006) admits as much, as we can see in the latter part of that quote. But what if we have no independent reasons for believing that Napoleon lost Waterloo? Well, in that case, our belief would've been unjustified anyway. At best, all the knowledge about the pill would have done is to have pointed out that our belief "was never justified in the first place". But if so, the argument seems very weak; indeed, Joyce (2006) and Kahane (2011) share this weakness in common. To recapitulate: EDAs are limited to contexts in which the *only* causes of S's belief that p are off-track processes; that is, if no epistemically respectable reasons are involved in the maintenance of S's belief that p. Furthermore, EDAs do not seem to be any better than simply asking S to provide reasons for her belief that p, and when she fails to do so, pointing out that her belief is not justified. The off-track process X and, by extension, EDAs just do not seem to be doing much work.

To avoid the genetic fallacy, we should examine *reasons*, not *causes*. If S have good reasons to believe p, the fact that S's belief that p is caused by an off-track process X is neither here nor there. If S does not have good reasons to believe p, she ought not believe p, regardless of the causes of her belief that p. If the claim that an off-track process X explains S's believing p is just a roundabout way of saying that S has no epistemically respectable reasons for believing p, then one wonders why the debunker does not just say that plainly, without dressing the argument up in trendy evolutionary terms.

EDAs and promiscuous evolutionary scepticism

The foregoing discussion assumes that it makes sense to talk about good reasons for believing particular propositions to be true, at least under a correspondence theory of truth and a realist semantics of those propositions. However, there are varieties of EDAs that might lead to *global* or *promiscuous* evolutionary scepticism: that is, EDAs whose epistemic premises assert that all our cognitive and perceptual faculties are off-track (or that we have no good reasons to believe that our cognitive and perceptual faculties are truth-tracking). Alvin Plantinga's (1993, 2011) Evolutionary Argument Against Naturalism (EAAN) is perhaps the most well-known of this kind of evolutionary scepticism (see Beilby 2002 for extended discussion). Plantinga's (2011) argument centres around whether or not our cognitive faculties—broadly con-

Author's personal copy

strued as "those faculties, or powers, or processes that produce beliefs or knowledge in us" (p. 311), including our perceptual faculties, memory, and *a priori* intuition among others—are reliable; in brief, Plantinga (2011) argues that, given a naturalistic account of the evolution of our cognitive faculties, it is unlikely that they are reliable (or, rather, we cannot assume that they are reliable). If so, then the multifarious deliberations of our cognitive faculties are thereby undermined. Plantinga (1993, 2011) famously uses this argument specifically against naturalism, but that need not detain us here. We shall look at this global evolutionary scepticism more generally, as well as one strategy to free some (but not all) kinds of beliefs from it. A global evolutionary debunking argument may be stated as follows, again following Kahane's (2011) scheme:

> *Causal premise*. All of S's beliefs are explained by X.
> *Epistemic premise*. X is an off-track process.
>
> [where an "off-track process" is just one that does not track truth]
>
> Therefore
> All of S's belief are unjustified

There are various strategies by which one might attempt to escape from such evolutionary scepticism, though a thorough evaluation of the EAAN and rejoinders to it goes far beyond the scope of this paper. However, one way out of this quagmire is to deny the epistemic premise, at least for certain *types* of beliefs. Take, for example, our commonsense beliefs, or beliefs about the medium-sized objects with which we usually interact. Even though there is no necessary connection between the deliverances of our cognitive faculties about medium-sized objects—adaptive behaviours could have evolved without true beliefs; indeed, adaptive behaviours could have evolved with wildly false beliefs—we might nevertheless argue that the best explanation for why our cognitive faculties produce the beliefs that they do is that these beliefs are *true* or approximately true. Indeed, Griffiths and Wilkins (2013) assert, regarding commonsense beliefs, that "it is hard to see what the basic evolutionary function of cognition could be other than tracking truth". So, one way to save a certain kind of belief from evolutionary scepticism is to argue from their pragmatic or evolutionary success to their truth; this kind of argument, Griffiths and Wilkins (2013) call a *Milvian Bridge*, after The Battle of Milvian Bridge, for which, according to some legends, the Emperor Constantine attributed his success to the truth of his Christian beliefs. They formulate the Milvian Bridge principle as follows:

> Milvian Bridge: *X* facts are related to the evolutionary success of *X* beliefs in such a way that it is reasonable to accept and act on *X* beliefs produced by our evolved cognitive faculties

Having built this so-called *Milvian Bridge* from the evolutionary success of commonsense beliefs to the truth of such beliefs, Griffiths and Wilkins (2013) proceed to argue that an *indirect* Milvian Bridge can be built to secure scientific beliefs, but not moral or religious beliefs.

Author's personal copy

Whether or not a Milvian Bridge can be built for commonsense beliefs without begging the question,[3] it is clear that no *direct* Milvian Bridge can be built for scientific beliefs. To wit, according to Griffiths and Wilkins (2013):

> Our cognitive faculties were selected because they tracked truth in the human *Umwelt*, not for their ability to do calculus, or to track truth about superpositions of particles at the quantum level. There is no direct Milvian bridge linking these particular cognitive processes to evolutionary success.

Just so. However, they contend that an *indirect* Milvian Bridge may be constructed to save our scientific beliefs from drowning in evolutionary scepticism:

> The reasons we have to think that our scientific conclusions are correct and that the methods we use to reach them are reliable are simply the data and arguments which scientists give for their conclusions, and for their methodological innovations. Ultimately, these have to stand up to the same commonsense scrutiny as any other addition to our beliefs. Thus, if evolution does not undermine our trust in our cognitive faculties, neither should it undermine our trust in our ability to use those faculties to debug themselves – to identify their own limitations, as in perceptual illusions or common errors in intuitive reasoning. Nor should it undermine our confidence in adopting new concepts and methods which have not themselves been shaped by the evolution of the mind, but whose introduction can be justified using our evolved cognitive faculties.

What Griffiths and Wilkins are saying here is that scientific beliefs and scientific methods can be assessed and justified in the court of commonsense thinking. This is enough to provide an indirect Milvian Bridge to save our scientific beliefs from evolutionary debunking.

Now, if this is all that it takes to provide an indirect Milvian Bridge, one wonders whether why such a bridge could not be constructed for other kinds of beliefs, say, moral or religious beliefs. In response regarding religious beliefs, Griffiths and Wilkins (2013) state, quite rightly, that "none of the leading accounts of the evolution of religious beliefs makes any reference to the truth or falsity of those beliefs when explaining their effects on reproductive fitness". But of course, the same can be said for our scientific beliefs, and as in the case of scientific beliefs, this simply rules out a *direct* Milvian Bridge, not an indirect one. It seems that if scientific beliefs can be saved from evolutionary debunking by resorting to the role commonsense beliefs (and the relevant cognitive faculties) in "debugging" them, at least some forms of religious reasoning can be saved as well. After all, it is not as though philosophical arguments about theism and atheism are based on radically different cognitive faculties than the ones that are meant to debug our beliefs, scientific, metaphysical, or otherwise. Indeed, the leading accounts of the evolution of religious beliefs explicitly assert that religious beliefs are based on normal cognition.

[3] After all, Griffith and Wilkins's (2013) inference to the best explanation is itself in the dock under this radical evolutionary scepticism.

Author's personal copy

As with most ideas, this suggestion is hardly new. In his response to the EAAN, Richard Swinburne argues for the conclusion that beliefs about the "metaphysical" are judged correct or incorrect on the basis of the same inductive criteria that apply to the "mundane". That is, both reasoning about commonsense and non-commonsense beliefs (e.g., beliefs about unobservable quarks, beliefs about unobservable gods) involve the assessing of evidence along the same or similar criteria. Swinburne (2004, pp. 353–354) writes:

> [T]here is no sharp dividing line between beliefs about the mundane and beliefs about the metaphysical. The latter are merely beliefs at the end of a spectrum of beliefs about larger and deeper matters. There is, we have seen, a selective advantage in having mechanisms that make inferences on mundane matters in accord with correct criteria of what is evidence for what. The higher animals, as well as humans, can make the predictions licensed by simple generalizations from observable data of particular interest to them. ... But humans are capable of more sophisticated reasoning; and they have a selective advantage in being able to add to their stock of true beliefs by reflection and experiment. The continuity of subject matter between the mundane and the metaphysical will lead us to use the same criteria in metaphysics.

The point is well made: if our commonsense belief-forming mechanisms are reliable, so are the belief-forming mechanisms they support and "debug". Swinburne accordingly concludes that although natural selection does not directly select for correct metaphysical beliefs, it indirectly supports the enterprise of arriving at such beliefs. Of course, we need to be extremely careful and reflective when we assess beliefs and evidence about the non-commonsensical, unobservable, or metaphysical; after all, our belief-forming systems are here operating outside the domain in which they evolved for their truth-tracking capabilities. Organisms with false commonsense beliefs are likely to be eliminated rather quickly, whereas the same does not apply to false scientific, philosophical, or religious beliefs. This means that false beliefs in these domains are likely survive much longer than false commonsense beliefs, but as Swinburne points out, "humans have the correct inductive criteria to weed them out, and, when combined with a desire in general to hold true beliefs – without which humans would not survive for long – humans will have a tendency in the course of time to acquire true metaphysical beliefs" (p. 354).

Swinburne builds an indirect Milvian Bridge to *metaphysical* beliefs, broadly construed, not distinguishing different kinds of unobservable causal entities (e.g., quarks, God).[4] This is where evolutionary debunkers of religious beliefs might depart from Swinburne. But on what basis might they do so? How might one distinguish between different kinds of unobservable causal entities without begging the question? One way to do this might be to maintain that while scientific beliefs and the methods through which they are achieved usually pass the "debugging" process unscathed, religious beliefs do not. But, of course, this claim must be argued for, and at least Griffiths and

[4] Some classical theists may object that God is not an "unobservable causal entity", nor indeed an entity of any kind (cf. Davies 1985; Mascall 1943). The philosophical significance of this theological assertion is under-explored in contemporary attempts to put science and theism in conversation.

Author's personal copy

Wilkins (2013) do not do so: they give no reason whatsoever to think that religious beliefs could not be justified in the court of commonsense in much the same way that than scientific beliefs can. Indeed, they finally concede:

> [D]ebunking is not disproving. If there are independent reasons for religious belief, their cogency is not removed by the fact that religious beliefs have evolutionary explanations.

We agree. But this just means that an indirect Milvian Bridge *may* be built for religious—and, indeed, other metaphysical—beliefs too. If religious (or, for that matter, anti-religious) beliefs are backed up by good arguments, then there is indeed an indirect Milvian Bridge via commonsense beliefs. Whether or not such arguments are available is, of course, not the domain of evolutionary biology or cognitive science but of philosophy and theology. As exciting as they are, evolutionary explanations of beliefs about gods ought not distract from the philosophical work of debugging such beliefs.

So, there does not seem to be a route to evolutionarily debunking religious beliefs from global evolutionary scepticism; at least, there does not seem to be one that does not simultaneously lead to the debunking of non-commonsense scientific beliefs. If commonsense beliefs can be saved from evolutionary debunking, and if scientific beliefs can be indirectly saved too, there seems to be no reason to think that religious and other metaphysical beliefs are any different.

EDAs and reliabilist arguments for religious beliefs

So far, it seems that EDAs have limited value as direct arguments against theism. However, this is not to say they have no value whatsoever. As we have seen, S's belief that p is caused by an off-track process X is irrelevant if S has good independent reasons for p, and does no philosophical work if S has no such reasons for p. Nevertheless, the fact that X is an off-track process *is* relevant if S's belief that p is based on misplaced trust in X. That is, if S offers an *explanation* of her belief that p as a *justification* of her belief that p, then pointing out that her belief that p is in actual fact wholly explained by an off-track process *should* undermine this belief; it does so by denying the epistemic premise in arguments with the following structure:

> *Causal premise*. S's belief that p is explained by X
> *Epistemic premise*. X is truth-tracking process.
> Therefore
> S's belief that p is unjustified

So, EDAs knock out certain sources of evidence, poison certain epistemic wells; in some cases, the well an EDA knocks out is the only one that sustains the relevant belief and if so, this belief will be unjustified. But in other cases, the epistemic agent has multiple sources of evidence; indeed, in some cases, casual explanations of the belief might not feature as a justification of the belief at all. It seems, then, that the applicability of an EDA depends in large measure on the sources of evidence available to the epistemic agent in question. This, in turn, depends in part on the success or

Author's personal copy

failure of various arguments, which if we are concerned with theism, are the business of philosophers of religion; however, it also depends crucially on one's epistemological theory.

Generally speaking, reliabilists insists that beliefs are justified when they are products of trustworthy mechanisms, whereas internalists link justification to reasons to which the epistemic agent has access. Evolutionary debunking arguments might pose greater challenges for reliabilists and other externalists than they do to internalists, because, very roughly, reliabilists explicitly bank on the reliability of the belief-forming mechanisms. They are committed to the epistemic premise above, which EDAs are designed to deny. So, it would seem, EDAs affect internalist and externalist philosophers of religion differently.

An internalist natural theologian constructs arguments and reasons for theism that consists mostly of beliefs that are widely accepted, or are at least publicly accessible (i.e., do not depend on religious premises). Richard Swinburne (2004)—an obvious candidate in the internalist, and indeed evidentialist camp—argues from general features of our empirically observable cosmos (e.g., contingency of existence, fine-tuned laws of nature, the existence of human agents capable of morality and reasoning, etc.) to the existence of God.[5] As is well-known, the project is to demonstrate that such features of the world would be probable under theism, and improbable under atheism. Any given belief in God might be *causally explained* in terms of revelation, ordinary cognitive mechanisms or religious experiences, but it is *justified* in terms of generally accessible arguments and reasons. For a natural theologian of this kind, the fact that our unreliable cognitive mechanisms might play a considerable role in the emergence of theistic beliefs is unproblematic: the rationality of theism rests on arguments and reasons, not on the reliability of our cognitive mechanisms. Therefore, biological and cognitive explanations of religion make a difference to this kind of natural theology only to the extent that one can form an argument directly to atheism from their results. EDAs cannot fulfil this role.

Of course, not all share this internalist project in natural theology. Reformed Epistemologists (RE; e.g., Plantinga 2000), for example, claim, very roughly, that belief in God could be rational even without any evidence of the kind that the internalist natural theologian requires. Indeed, proponents of RE are often sceptical that such evidence could justify the belief in God in any meaningful way. Instead, belief in God is justified because it immediately presents itself to the believer. For justified belief in God, the argument goes, it is enough that the believer has counterarguments to defeaters (i.e.,

[5] This is not to deny that Swinburne thinks that religious experiences have a role to play in justifying religious belief. Indeed, Swinburne argues that "One who has had a religious experience apparently of God has, by the Principle of Credulity, good reason for believing that there is a God—other things being equal" (p. 324). But what if things are not equal? Swinburne concedes, as we do, that one who has had a religious experience but has other reasons to believe "that it is significantly more probable than not that there is no God" (p. 326) is *not* justified in believing that there is God. But what if it can be shown that "special considerations"—such as the influence of hallucinogens—render one's religious experience questionable? Swinburne concedes, as we do, that the religious experience would not provide sufficient reason for believing that there is a God. However, neither Swinburne nor we believe that everything rides on the religious experience. Even if the religious experience itself fails to provide sufficient reason for believing that there is a God, other arguments are available. The validity and soundness of these arguments are not affected by the reason-providing poverty of the religious experience.

Author's personal copy

arguments that directly undermine belief in God, either by showing that it came from an unreliable source or that theism is improbable or impossible). For reliabilists, the deliverances of our cognitive faculties are innocent until proven guilty. However, EDAs put these cognitive faculties in the dock, accusing them of unreliability, and thereby undermine the beliefs they produce (see Dawes and Jong 2012, for a recent EDA-type argument against Reformed Epistemology). If so, EDAs are much more threatening to Reformed Epistemologists than to evidentialists.[6] Rather than an argument against theism, perhaps EDAs are more fruitfully thought of as arguments against a popular trend in contemporary philosophy of religion. Interestingly, however, there have been various suggestions to the contrary, albeit vague ones, which contemporary cognitive and evolutionary theories of religious belief are consistent with theism precisely via Reformed Epistemology (e.g., Clark and Barrett 2011). Perhaps this is so, but proponents of such claims will have to wrestle with EDAs; we shall leave this task to them.

References

Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford: Oxford University Press.

Barrett, J. (2004). *Why would anyone believe in God?*. Walnut Creek: AltaMira Press.

Beilby, J. (Ed.). (2002). *Naturalism defeated?. Essays on Plantinga's evolutionary argument against naturalism*. Ithaca: Cornell University Press.

Benfey, O. T. (1958). August kekulé and the birth of the structural theory of organic chemistry in 1858. *Journal of Chemical Education*, *35I*, 21–23.

Bering, J. (2011). *The God instinct: The psychology of souls destiny and the meaning of life*. London: Nicholas Brealey Publishing.

BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy*, *5*, 53–73.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

Brosnan, K. (2011). Do the evolutionary origins of our moral beliefs undermine moral knowledge? *Biology and Philosophy*, *26*, 51–64.

Clark, K., & Barrett, J. (2011). Reidian religious epistemology and the cognitive science of religion. *The Journal of the American Academy of Religion*, *79*, 639–675.

Darwin, C. (1881). Letter 3230–Charles Darwin to William Graham 3 July 1881. Retrieved from http://www.darwinprojectacuk/entry-13230 on March 5 2012

Davies, B. (1985). *Thinking about God*. London: Geoffrey Chapman.

Dawes, G. W., & Jong, J. (2012). Defeating the christian's claim to warrant. *Philo*, *15*, 127–144.

De Cruz, H., Boudry, M., De Smedt, J., & Blancke, S. (2011). Evolutionary approaches to epistemic justification. *Dialectica*, *64*, 517–535.

Dennett, D. (2006). *Breaking the spell: Religion as a natural phenomenon*. New York: Penguin.

Goldman, A. I. (1986). *Epistemology and cognition*. Cambridge: Harvard University Press.

Goldman, A. I. (2008). Immediate justification and process reliabilism. In J. Smith (Ed.), *Epistemology: New essays* (pp. 63–82). Oxford: Oxford University Press.

Griffiths, P., & Wilkins, J. (2013). Evolutionary debunking arguments in three domains: Fact value and religion. In G. Dawes & J. Maclaurin (Eds.), *A new science of religion* (pp. 133–146). New York: Routledge.

Hoyningen-Huene, P. (2006). Context of discovery versus context of justification and thomas kuhn. *Archimedes*, *14*, 119–131.

[6] This is probably the reason why much of the EDA debate in philosophy of religion has so far been preoccupied with defending the truth-tracking nature of the cognitive mechanisms that (according to CSR) produce theistic beliefs. See, for example Clark and Barrett (2011), Murray (2009), and Murray and Goldberg (2009).

Author's personal copy

Hume, D. (2008). Dialogues and Natural history of religion. J. C. A. Gaskin (Ed.), Oxford:Oxford University Press. (Original work published 1757).
Jong, J. (2012). Explaining religion (away?): Theism and the cognitive science of religion. *Sophia*, *52*, 521–533.
Joyce, R. (2006). *The evolution of morality*. Cambridge: MIT Press.
Kahane, G. (2011). Evolutionary debunking arguments. *Noûs*, *45*, 103–125.
Klement, K. (2002). When is genetic reasoning not fallacious? *Argumentation*, *16*, 383–400.
Leech, D., & Visala, A. (2011). The cognitive science of religion: A modified theist response. *Religious Studies*, *47*, 301–316.
Leech, D., & Visala, A. (2012). How relevant is the cognitive science of religion to philosophy of religion. In: Y. Nagasawa (Ed.), *Scientific approaches to the philosophy of religion*. Basingstoke:Palgrave.
Mascall, E. L. (1943). *He who is*. London: Darton, Longman & Todd.
Murray, M. (2009). Scientific explanations of religion and the justification of religious belief. In M. J. Murray & J. Schloss (Eds.), *The believing primate: Scientific philosophical and theological reflections on the origin of religion* (pp. 168–177). Oxford: Oxford University Press.
Murray, M., & Goldberg, A. (2009). Evolutionary accounts of religion: Explaining and explaining. In M. J. Murray & J. Schloss (Eds.), *The believing primate: Scientific philosophical and theological reflections on the origin of religion* (pp. 179–199). Oxford: Oxford University Press.
Plantinga, A. (1993). *Warrant and proper function*. Oxford: Oxford University Press.
Nozick, R. (1981). *Philosophical explanations*. Cambridge: Harvard University Press.
Plantinga, A. (2000). *Warranted Christian belief*. Oxford: Oxford University Press.
Plantinga, A. (2011). *Where the conflict really lies?*. Oxford: Science religion and naturalism Oxford University Press.
Reichenbach, H. (1938). *Experience and prediction*. Chicago: University of Chicago Press.
Ruse, M . (2006). Evolutionary ethics and contemporary biology. In G. Boniolo & G. de Anna (Eds.), *Is Darwinian metaethics possible (and if it is is it well-taken)?* (pp. 13–26). Cambridge: Cambridge University Press.
Ruse, M. (2009). Evolution and ethics: The sociobiological approach. In M. Ruse (Ed.), *Philosophy after Darwin* (pp. 489–511). Princeton: Princeton University Press.
Salmon, W. (1970). Bayes's theorem and the history of science. In R. Stuewer (Ed.), *Historical and philosophical perspectives of science* (pp. 68–85). Minneapolis: University of Minnesota Press.
Salmon, W. (1984). *Logic*. Englewood Cliffs: Prentice-Hall.
Siegel, H. (1980). Justification discovery and the naturalizing of epistemology. *Philosophy of Science*, *47*, 297–32.
Swinburne, R. (2004). *The existence of God* (2nd ed.). Oxford: Oxford University Press.
Street, S. (2006). A darwinian dilemma for realist theories of value. *Philosophical Studies*, *127*, 109–166.
Tremlin, T. (2006). *Minds and gods: The cognitive foundations of religion*. Oxford: Oxford University Press.
Wilson, D. S. (2002). *Darwin's cathedral: Evolution religion and the nature of society*. Chicago: University of Chicago Press.