Darwinism in Metaethics: What If the Universal Acid Cannot Be Contained?

Eleonora Severini *

eleonora.severini@uniroma1.it

Fabio Sterpetti *

fabio.sterpetti@uniroma1.it

ABSTRACT

The aim of this article is to explore the impact of Darwinism in metaethics and dispel some of the confusion surrounding it. While the prospects for a Darwinian metaethics appear to be improving, some underlying epistemological issues remain unclear. We will focus on the so-called Evolutionary Debunking Arguments (EDAs) which, when applied in metaethics, are defined as arguments that appeal to the evolutionary origins of moral beliefs so as to undermine their epistemic justification. The point is that an epistemic disanalogy can be identified in the debate on EDAs between moral beliefs and other kinds of beliefs, insofar as only the former are regarded as vulnerable to EDAs. First, we will analyze some significant debunking positions in metaethics in order to show that they do not provide adequate justification for such an epistemic disanalogy. Then, we will assess whether they can avoid the accusation of being epistemically incoherent by adopting the same evolutionary account for all kinds of beliefs. In other words, once it is argued that Darwinism has a corrosive impact on metaethics, what if its universal acid cannot be contained?

Keywords: Darwinism; Metaethics; Epistemic Justification; Evolutionary Debunking Arguments; Moral Realism; Scepticism.

---

*   Sapienza University of Rome, Department of Philosophy.

1. Introduction

"Some scientific thinkers, while not themselves philosophers, make philosophers necessary. Charles Darwin is an obvious case" (Hodge and Radick 2003, 1). The philosophical impact of Darwin's thought goes far beyond his field of research and affects not only philosophy of biology and epistemology, but also social sciences and moral philosophy (Lewens 2005). Furthermore, "[n]owadays, philosophical 'naturalism' pretty much means philosophy driven by mainly insights from Darwin" (Rosenberg 2016, 24). However, Darwinian approaches to moral philosophy have been widely disputed. The basic idea underlying these critiques is that "if we are the products of blind natural selection, then morality and value are merely reflections of our subjective attitudes, and that in that case everything is permitted and nothing matters" (Kahane 2011, 103). On this view, naturalizing moral philosophy by explaining morality in Darwinian terms will have some worrisome consequences for our standard conceptions of morality.

Recently, these worries have been summarized in the debate on the Evolutionary Debunking Arguments (EDAs). In our view, although these arguments have been employed in several moral domains, e.g. normative ethics and applied ethics, Darwinian theory is particularly pertinent in answering metaethical questions. There are various metaethical questions, but they share a concern for the nature of morality and, for instance, for the justification of our moral beliefs. Thus, when applied in metaethics, EDAs are defined as arguments that appeal to the evolutionary genealogy of our moral beliefs in order to undermine their epistemic justification (Kahane 2011). This might pave the way to some form of moral skepticism, i.e. the view according to which one cannot have moral knowledge or justified moral beliefs (Sinnott-Armstrong 2015).

The aim of this article is to show that, while the prospects for a Darwinian metaethics appear to be improving, some underlying epistemological issues remain obscure. In particular, we will focus on the epistemic disanalogy between moral beliefs and other kinds of beliefs (e.g. scientific, mathematical, etc.), insofar as it is usually only the former that are regarded as being vulnerable to EDAs. We will explore some significant debunking positions in metaethics in order to show that they do not provide adequate justifications for the epistemic disanalogy, and assess whether they can avoid the accusation of being epistemically incoherent.

More precisely, we will argue that, since there is no principled way to neatly distinguish between, on the one hand, domains whose beliefs can be epistemically justified by natural selection and, on the other, domains whose beliefs cannot be epistemically justified by

natural selection, the debunker risks being epistemically incoherent in claiming that some domains can be justified on the basis of natural selection, while certain other domains cannot be justified in the same way. If the debunker wishes to avoid the epistemic incoherence objection, she has two options: either (a) she adopts a realist attitude toward every domain, or (b) she adopts a 'deflationary' attitude toward every domain. As we will try to make clear, if the debunker opts for (a), she risks being compelled to adopt a realist stance also toward some domains which are not accessible to science, i.e. domains whose beliefs the debunkers usually try to debunk precisely because those beliefs cannot be accounted for in naturalist terms (e.g. morality).[1] Thus, the debunker risks being unable to claim that her position is a naturalist position. If, on the contrary, the debunker chooses option (b), i.e. she adopts a 'deflationary' attitude toward every domain, then she is certainly able to claim that her position is a naturalist position. But this comes at a price for the debunker, namely renouncing a realist attitude toward the beliefs of certain domains which are usually realistically interpreted by the majority of the naturalists. This 'revisionary' attitude may be unpalatable to some debunkers. Nevertheless, given that naturalism is the most fundamental element for a debunker position, we will suggest that (b) is the more promising option for the debunker. Finally, we will argue that (b) does not necessarily lead to skepticism.

So, the question we will try to answer is the following: Once it is argued that Darwinism has a corrosive impact on metaethics, what if its universal acid cannot be contained?[2]

---

[1] Here, some clarifications about the relationship between realism and naturalism are in order. Even though there are many forms that realism and naturalism can take, for the purpose of this paper, realism can be understood as the view that certain entities (e.g. moral facts) exist independently of anyone's beliefs, and our beliefs about those entities are true insofar as they correspond to those entities. Naturalism can be understood as the view that philosophers should refute any entity that cannot be empirically investigated, and our beliefs are true insofar as they can be empirically supported. It should be noted that realism and naturalism do not overlap. While some try to combine realism and naturalism (e.g. Railton 1986; Brink 1989), the majority of moral philosophers adopt either a non-naturalist realist view (e.g. Shafer-Landau 2003; Parfit 2011) or a naturalist non-realist view (e.g. Gibbard 1990; Street 2006). The main difficulty in combining realism and naturalism concerns how to conceive of beliefs about *a priori* domains, i.e. domains that are *in principle* empirically inaccessible (see sections 3.3 and 4).

[2] The well-known metaphor of Darwinism as a universal acid eroding every traditional concept, morality included, comes from Daniel Dennett: "Darwin's idea is a universal solvent, capable of cutting right to the heart of everything in sight. The question is what does it leave behind?" (Dennett 1995, 521). Dennett's metaphor was the inspiration behind the title of Rosenberg's article "Darwinism as Philosophy: Can the Universal Acid Be Contained?" (2016), from which the title of our paper is derived.

2. *The Evolutionary Debunking Arguments in Metaethics*

A considerable part of the current debate in metaethics is devoted to handling the evolutionary explanation of morality, and EDAs in metaethics are flourishing. Starting from the evolutionary origins of our moral beliefs, EDAs seek to undermine the epistemic standing of such beliefs[3] (Kahane 2011). More precisely, when proposed in metaethics, EDAs are intended to undermine the realist view of morality (Street 2006; Joyce 2006; Kitcher 2011).

According to moral realism, there are moral facts that exist insofar as they are "stance-independent", i.e. they would exist independently of any particular perspective (Shafer-Landau 2003), and moral beliefs are true insofar as they accurately depict those moral facts. Thus, moral truths can be construed as being determined by independent moral facts.[4]

In this regard, Darwinism provides us with a competing account for morality, according to which our moral attitudes are grounded in evolutionary origins. Thus, our propensity to have moral beliefs, i.e. roughly speaking, beliefs about right and wrong, is a biological adaptation.[5] In other words, moral beliefs have been promoted by natural selection not because they were true, but because they were advantageous, i.e. they emerged through natural selection because this way of thinking provided our ancestors with some sort of reproductive advantage (Ruse and Wilson 1986). The problems that arise for the epistemic status of moral beliefs in light of Darwinism have been clearly pointed out by Sharon Street in her "Darwinian Dilemma" (DD) (Street 2006). The DD tries to spell out the relation between the evolutionary influences on our moral beliefs, on the one hand, and the independent moral truths that moral realism posits, on the other.

The starting point of the DD is the Darwinian assumption that evolutionary "forces have played a tremendous role in shaping the content of human evaluative attitudes" (Street

---

[3] It is important to specify that, though there are various ways to understand an epistemic appraisal, we are here concerned with the *epistemic* justification of moral beliefs, i.e. whether "moral beliefs and claims can be justified or known or shown to be true" (Sinnott-Armstrong 1996, 5).

[4] Some authors refer the 'independence requirement' to 'moral facts', while others speak simply of 'independent moral truths'. Since the distinction between 'moral truths' and 'moral facts' is irrelevant for the purposes of this article, we will use these expressions interchangeably.

[5] Whether morality is an adaptation or a byproduct is still an open question in the ongoing debate involving both biologists and moral philosophers. However, although the debunkers generally refer to an adaptationist account of morality, they are not necessarily committed to such an account. As Kahane underlines, it is "important to see that it does not matter here whether any particular evolutionary explanation is true. What matters is that some such story is likely to be true" (Kahane 2011, 111).

2006, 109). The challenge for the moral realist is to explain the relation between these evolutionary influences on our evaluative attitudes and the independent truths that realism posits. Street argues that the realist has two options, which assume the form of a dilemma: she may either (1) deny that there is a relation between evolutionary influences on our evaluative attitudes and the independent moral truths, or (2) assert such relation. Street shows that the realist cannot give a satisfactory account of this relation whichever horn of the DD she embraces.

On the first horn of the DD, the realist may claim that there is no relation between evolutionary influences on our moral beliefs and independent moral truths. In this case, evolutionary forces must be viewed as distorting influences on our moral attitudes, having shaped us in ways that have nothing whatsoever to do with moral truths. If we take this claim and combine it with the DD's main assumption that our evaluative judgments have been shaped by evolutionary forces, then we are left with the skeptical conclusion that our moral beliefs are mostly off track. This conclusion is clearly unacceptable for the moral realist.

On the second horn of the DD, the realist may claim that there is a relation between evolutionary influences and independent moral truths, namely that natural selection has favored those ancestors who were able to grasp those truths. On this view, the capacity to track independent moral facts proved advantageous for the purposes of survival and reproduction. But this account, according to Street, is empirically untenable, i.e. it is refuted by current science. The details of Street's argumentation as to why such an account would be unacceptable on scientific grounds are not relevant here. Instead, it is worth noting that, contrary to Street's conclusion, the second horn of the DD offers the most tempting route for a moral realist to take in order to avoid the moral skepticism of the first horn.

According to Street, there are two competing explanations for the emergence of an adaptive capacity: the Tracking Account (TA) and the Adaptive Link Account (ALA). According to TA, the relation between our beliefs and the domain to which those beliefs refer is a tracking relation, i.e. there is a sort of 'representational correspondence' between the content of our beliefs and that portion of the (external) world to which those beliefs refer. In other words, in order to make TA plausible, one has to assume that the capacity to form and use (some kind of) beliefs has a survival value insofar as the beliefs formed are true (see, e.g. Millikan 1984). In contrast, according to ALA, the relation between our beliefs and the domain to which those beliefs refer is not a relation of tracking, but rather one driven by the adaptive advantage that some responses to some circumstances may afford us, independently from the correspondence between the content of our beliefs and the state of affairs in the (external) world.

Thus, the realist can try to embrace the second horn of Street's DD, and assert a truth-tracking relation between evolutionary forces and moral truths. That is to say, the realist

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

may claim that TA is the right way to account for the relation between evolutionary influences on our moral beliefs and independent moral truths. Thus, TA's core claim about morality is to say that "making […] such evaluative judgements contributed to reproductive success because they are true, and it proved advantageous to grasp evaluative truths" (Street 2006, 128). This is precisely the strategy adopted by some critics of EDAs (e.g. Enoch 2010; Skarsaune 2011; Graber 2012; Artiga 2015). These authors argue, in various ways, that the evolutionary explanation of morality, though correct, does not debunk anything, because the evolutionary process is not entirely insensitive to external moral truths. These accounts are committed to the view according to which natural selection *does care* about moral truths,[6] i.e. moral beliefs are reproductively advantageous because they are true, and they are true because they 'track' some independent moral truths.

Contrary to this view, as noted above, Street argues that the right way to give an evolutionary explanation of the content of our moral beliefs is ALA, not TA. In this perspective, "making such [evaluative] judgements contributed to reproductive success not because they were true or false, but rather because they got our ancestors to respond to their circumstances with behavior that itself promoted reproductive success […]" (Street 2006, 128-129). Asserting that the primary function of some set of beliefs is to promote a certain kind of adaptive behavior means that the tracking relation between the content of those beliefs and the external world is not a necessary requirement for those beliefs to be adaptive.[7] According to Street, moral beliefs are beliefs of this kind, given that "[t]he best causal accounts of our evaluative judgements, whether Darwinian or otherwise, make no reference to the realist's independent evaluative truths" (Street 2006, 155).

Now the question is: How can we determine whether TA or ALA provides the right evolutionary account for moral beliefs? This is a crucial question and it can be framed in more general terms. As Helen De Cruz and her co-authors note, "evolutionary considerations are being used to either justify or debunk a wide variety of beliefs" (De Cruz *et al.* 2011, 518). As such, there is no single way of drawing straightforward implications from evolutionary considerations for the epistemic standing of our beliefs. Following De

---

[6] On the relation between natural selection and truth, see Boudry and Vlerick (2014).

[7] The point is that even some false beliefs may be adaptive. If a tracking relation were necessary for a belief to be adaptive, at least *some* false beliefs (i.e. those which do not track *any* aspect of reality) could not have been adaptive. But adaptive misbeliefs are routinely produced because they result in some reproductive advantage, even if they do not track any *real* aspect of reality. Think of supernatural beliefs: "If belief in an omniscient, omnipotent deity is adaptive because it inhibits detectable selfish behaviour […], this will be the case whether or not such a being actually exists. If such a being does not exist, then we have adaptive misbelief" (McKay and Dennett 2009, 507). Thus, the presence of a tracking relation cannot be regarded as a necessary requirement for a belief to be adaptive. On adaptive misbeliefs, see also De Cruz *et al.* 2011.

Cruz *et al.* (2011), indeed, we can recognize two distinct evolutionary approaches to epistemic justification: the first approach relies on Evolutionary Arguments (EAs), while the second relies on the aforementioned EDAs.

The first position, supported by EAs, "contends that natural selection will tend to pick out and propagate those types of beliefs and judgments that correspond with the state of the world" (De Cruz *et al.* 2011, 518). In view of this, the supporters of EAs conclude that our evolved beliefs are justified and true. Thus, if we opt for an EA for morality, then we can conclude that our moral beliefs are justified and true. Moreover, if we recall TA, according to which our beliefs contributed to reproductive success because they are true, we should note that TA is equivalent to an EA.

The second position relies on EDAs. An EDA is constructed by negating at least one of the EA's crucial premises, "in particular about the relative importance of natural selection, and about its truth-tracking ability" (De Cruz *et al.* 2011, 518). In light of this, the debunkers conclude that our beliefs, while useful, are off track and hence are not justified. Thus, if we opt for an EDA for morality, then we can conclude that our moral beliefs are not justified and cannot be taken to be true. If we recall ALA, according to which our beliefs contributed to reproductive success not because they were true, but rather because they prompted our ancestors to respond to certain circumstances in ways that were adaptive, we can see that ALA is equivalent to an EDA.

Whether one should support EAs or EDAs in morality is beyond the scope of this paper. Thus, we do not enter into such a debate. However, having described the distinction between EAs and EDAs, and having presented the DD, we will deploy this distinction and the DD so as to illustrate an epistemological challenge that can be raised against EDAs' supporters.

3. The Epistemic Incoherence Objection

The epistemological challenge that can be raised against EDAs' supporters is the epistemic incoherence objection.[8] As we will try to make clear below, the epistemic incoherence objection deals with the difficulty of clearly demarcating those domains whose beliefs have to be accounted for in terms of TA and those whose beliefs have to be accounted for in terms of ALA. As stated above, supporting TA in certain domains (for instance, morality) would allow one to maintain a realist position in such domains (for instance, realism in metaethics). After having sketched how this objection is usually construed (3.1), we will consider a different way of construing it (3.2). Then, we will underline how the strategy

---

[8]   For other epistemological difficulties facing EDAs' supporters, see De Cruz *et al*. 2011.

one adopts in confronting the epistemic incoherence objection may affect one's commitment to naturalism (3.3).

3.1. The Epistemic Incoherence Objection as Usually Construed

In order to present the epistemic incoherence objection, let us turn once again to the DD. Given the distinction between EAs and EDAs, we saw that Street supports an EDA in metaethics. In other words, Street argues, against moral realism, that morality has to be accounted for in terms of ALA rather than TA. As we have seen above, according to ALA, "as a result of natural selection, there are in living organisms all kinds of mechanisms that serve to link an organism's circumstances with its responses in ways that tend to promote survival" (Street 2006, 127). Let us now focus on how Street explains how ALA works:

ALA. A straightforward example of such a mechanism is the automatic reflex response that causes one's hand to withdraw from a hot surface […]. Such mechanisms serve to link certain kinds of circumstances – the presence of a hot surface […] – with adaptive responses – the immediate withdrawal of one's hand […]. (Street 2006, 127).

In an analogous way, Street holds that "the function of evaluative judgements from an evolutionary point of view is not to 'track' independent evaluative truths, but rather to get us to respond to our circumstances in ways that are adaptive" (Street 2006, 157-158). Therefore, Street supports an antirealist stance in metaethics, according to which our moral beliefs do not depict any independent moral reality, but instead constitute adaptive responses to certain circumstances.

The problem for such a debunker position derives from the fact that while Street refers to ALA to explain moral beliefs, she refers to TA to explain beliefs in domains besides morality. Street explicitly claims that:

the Darwinian Dilemma doesn't go through against realism about non-evaluative facts such as facts about fires, predators, cliffs, and so on. In short, the difference is that in the case of such non-evaluative facts, unlike in the case of evaluative facts, the tracking account prevails as the best explanation of our capacity to make the relevant sort of judgement. In order to explain why it proved advantageous to form judgements about the presence of fires, predators, and cliffs, one will need to posit in one's best explanation that there *were indeed* fires, predators, and cliffs, which it proved quite useful to be aware of […]. (Street 2006, 160, fn. 35).

In other words, Street embraces an EDA for morality and an EA for other domains, e.g. scientific beliefs, common sense beliefs or beliefs based on perceptions. In so doing, Street's position is epistemically incoherent. Indeed, it could be claimed that once we have accepted an evolutionary account of the epistemic justification of our beliefs, then it is

epistemically incoherent to support antirealism in one domain and realism in other domains. The problem, critics say, is that if one goes evolutionary, it is difficult to contain 'the corrosive acid' of Darwinism to a specific set of beliefs. Thus, the debunker's epistemic incoherence may be, *prima facie*, construed as indicating that she is on a slippery slope: if one adopts a debunking position in metaethics, then, to be coherent, she risks embracing a full-blooded skeptical position, i.e. a skeptical position pertaining to all kinds of belief. If we claim that our belief-forming processes are *evolved* cognitive mechanisms, and we also maintain that evolution is not always or for the most part a truth-tracking process, then it is difficult to discriminate between those sets of beliefs which reliably track the truth from those sets of beliefs which are unreliable. By way of clarification, Vavova (2015) offers a comparison between moral beliefs and perception:

> Do we have good reason to think that perception would lead us to true beliefs about our surroundings? Not if 'good' reason is understood as an appropriately independent reason: for if we set aside all that is in question, we must set aside all beliefs gained by perception. This includes all scientific beliefs, like the belief that evolutionary theory is true. Without those, we cannot evaluate the rationality of beliefs formed by perception. (Vavova 2015, 83).

So, the argument goes, if we deny that our evolved mechanisms lead us to (produce mostly) true beliefs, there is the threat of skepticism. But the idea of supporting such a *global* skepticism is surely beyond the debunkers' intentions. Indeed, if EDAs imply a sort of global skepticism, the untenability of moral realism, which is the claim that EDAs usually support, is undermined as much as the realist's claim. In other words, if *nothing* can be justified because evolutionary influences prevent us from justifying *any* claim, then it is surely not only moral realism that is in trouble.[9]

However, to face this formulation of the epistemic incoherence objection, one may consider the task of evaluating the reliability of the beliefs produced in each domain as an *empirical* question. It is science that investigates which sets of beliefs reliably track features of reality. Despite Vavova's above claim, this seems a viable route for the debunker. Indeed, Vavova seems to assume that we have to endorse some sort of absolute view on justification, according to which if we are not able to show that the justification of our beliefs is independent from any domain whatsoever, we cannot really claim to be able to justify our beliefs.

But there is no reason why this claim should be assumed by the debunkers, at least by those who intend to maintain a naturalist stance, in the very general sense of taking into account how science proceeds in their mode of philosophizing. Indeed, it seems that in

---

9   On the issue of whether EDAs are self-defeating arguments, see Sterpetti (2015), and for a reply, Kyriacou 2016.

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

order to adjudicate between the justifications of beliefs that we form in some domain, we can safely proceed in the same manner in which science usually proceeds.[10]

In this perspective, it is not clear why the independence requirement – which Vavova correctly claims is necessary in order to assess whether we have good reasons to rationally maintain some set of beliefs – should be understood in such a demanding way. The idea that we have to set aside "all that is in question", in order to avoid begging the question, is usually construed in the scientific inquiry as merely implying that we have to avoid assuming the truth of those beliefs that belong to the very domain we are currently investigating. The independence requirement does not amount to assuming that we have to avoid relying on any sort of beliefs, including those beliefs that are only indirectly related (or irrelevant) to the domain currently under investigation, calling everything into question at the same time (Vlerick and Broadbent 2015).

Consider Vavova's own example. Surely there is a sense in which it is correct to say that our perception *does not* lead us to true beliefs about our surroundings, since we know that our perception may well deceive us. We know this because we can confront our perceptions with the results of reliable scientific findings. And the reliability of those scientific findings does not directly and relevantly depend on the reliability of our perception. And certainly, to accept this claim does not amount to embracing skepticism. Moreover, if in order to avoid skepticism, we claim that perception cannot lead us to false beliefs about our surroundings, then we should be prepared to support the claim that it is impossible for us to have false perceptions. In other words, we should defend some kind of epistemic infallibilism, which may be deemed a legitimate epistemological position, but it is usually regarded as incompatible with a naturalist perspective, given that naturalism is usually thought to imply a commitment to some (even minimal) form of fallibilism (Stump 1991).

Vavova tries to put pressure on the debunker by using such a demanding constraint on what may be considered as an "appropriately independent reason" and on what has to be left aside when providing epistemic justification for any given belief. Indeed, a crucial point in her defense of moral realism is the following: if, following the debunkers, we assume that in justifying our moral attitude we have to put aside *any* moral assumption in order not

---

[10] It is important to note that we are assuming here that the debunker and her opponent adopt a naturalist stance, because EDAs are problematic only if one commits oneself to some form of naturalism. If, instead, one adopts an anti-naturalist stance, then one may well ignore the challenges that arise from EDAs (see, e.g. Bogardus 2016). As to how to understand 'naturalism', see fn. 1. We are not concerned here with any specific view of naturalism, nor will we survey the many criticisms that have been raised against this (family of) view(s) (for a survey, see Clark 2016; Papineau 2016; Rosenberg 1996).

to beg the question, then we can indeed have no idea of what a moral belief is like. If we put aside *any* moral assumption, we cannot determine whether or not any given belief is a moral belief. But this means that "we cannot determine if we are likely to be mistaken about morality if we can make no assumptions at all about what morality is like" (Vavova 2015, 92). The debunker, so the argument goes, will thus be unable to support her chief claim, i.e. that "true evaluative beliefs and the adaptive evaluative beliefs come apart" (Ibid.), given that she cannot compare the content of some set of adaptive evaluative beliefs (A) with the content of some set of moral beliefs (B) determined independently of (A), and prove that the content of A-beliefs and the content of B-beliefs diverge. But if (A) and (B) cannot be compared and proved to be divergent, then it is not possible for the debunker to claim that evolution influenced our evaluative belief-forming process in such a way that it is unreasonable to think that the content of our adaptive evaluative beliefs corresponds to the content of independent moral truths as posited by the moral realists.

But this way of representing what the debunkers claim or aim to show is quite unfair. The debunker's position may be more fairly formulated as follows: moral realists *do* claim that moral beliefs are the way they are because they correspond to independent moral facts, i.e. facts which are causally disconnected from the empirical world. The debunker may well grant, for argument's sake, that realists do know the content of moral beliefs. The problem is not in determining the content of moral beliefs, as stated by moral realists, in order to check whether it matches the content of adaptive beliefs, as stated by moral debunkers; rather, the problem is how to support the realist claim that the content of moral beliefs is determined by a tracking relation, if moral truths are independent from the empirical world. The issue is that the realist should be able to demonstrate that our cognitive mechanisms can track moral truths, given: (1) the independence of moral truths; and (2) the influences that evolution may have had on the formation of such cognitive mechanisms. In other words, it is the realist that should give us a plausible evolutionary account of how we track independent moral facts, since the independence claim is a characteristic feature of moral realism. On the other hand, the debunker may legitimately *a posteriori* confront the moral realist account of the origin of our moral beliefs with findings that could be relevant in order to appraise the empirical plausibility of such a realist account.

Now, we will put aside this kind of construal of the epistemic incoherence objection, and, for the sake of argument, in what follows we will assume that different domains may well be evaluated independently from each other, so that global skepticism does not necessarily follow if we doubt the reliability of our cognitive faculties in some domains on evolutionary grounds.

3.2.   *The Epistemic Incoherence Objection, Reloaded*

The difficulty that we would like to address pertains to how the sets of beliefs that require TA are distinguished from the sets of beliefs that require ALA by the debunkers themselves. More precisely, we will try to assess whether a *principled distinction* between those two accounts is well supported by the examples given by the debunkers to illustrate such a distinction.

To this end, it will be useful to recall the example of TA given by Street for domains other than morality:

TA.   First consider truths about a creature's manifest surroundings – for example, that there is a fire raging in front of it, or a predator rushing toward it. It is perfectly clear why it tends to promote reproductive success for a creature to grasp such truths: the fire might burn it to a crisp; the predator might eat it up. (Street 2006, 130).

Now, recall the example given by Street to illustrate ALA, cited previously:

ALA.   A straightforward example of such a mechanism is the automatic reflex response that causes one's hand to withdraw from a hot surface […]. Such mechanisms serve to link certain kinds of circumstances – the presence of a hot surface […] – with adaptive responses – the immediate withdrawal of one's hand […]. (Street 2006, 127).

In response to both passages, one can ask: Is there really a relevant and clear difference between the case of a hot surface, reported as an example of ALA, and those of a fire and a predator, reported as examples of TA?

Recall that, for Street, the distinction between ALA and TA lies in the fact that in order to *explain* why it proved advantageous to form judgments about non-evaluative facts, e.g. "the presence of fires, predators, and cliffs," one has to posit "in one's best explanation that there *were indeed* fires, predators, and cliffs" (Street 2006, 160, fn. 35).

The difference between ALA and TA should indeed be due to the irrelevance of truth-tracking considerations for the former and the relevance of truth-tracking considerations for the latter. But Street's proposed cases seem not to support such a principled distinction between the two kinds of accounts, i.e. TA and ALA. Indeed, even in the case that is intended to illustrate ALA, truth-tracking considerations seem to intervene in the best explanation of why having a certain response proved to be adaptive.

If an evolved cognitive mechanism is built in such a way that the correct adaptive response – "the immediate withdrawal of one's hand" – has to be given by the organism if certain circumstances obtain, then it is relevant that the beliefs produced by such mechanism are able to track the truth of at least *some* aspect of the current state of affairs of the world in order to ascertain whether or not those circumstances actually obtain (e.g. whether or not a "hot surface" is actually present in the organism's environment). So, it

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

seems that in order to *explain* why it proved advantageous to withdraw one's hand, one needs to posit in one's best explanation that there *was indeed* a hot surface.

The problem for the debunker is that if, in an example of ALA relative to a domain other than morality, truth considerations seem to be relevant, it is difficult to see why ALA relative to morality should be insensitive to truth considerations. It may be objected that in the case of the hot surface, even if there is a sort of representation of the actual presence of the hot surface in the belief, this is not the salient feature that has to be considered in order to account for the adaptiveness of the response triggered by the belief. The problem in such a strategy is that if the presence of some truth-tracking considerations is also permitted for ALA-domains, then the very distinction between ALA and TA seems to fade away. Indeed, if the presence of some capacity of reliably detecting and representing some aspect of the state of affairs of the actual environment (i.e. tracking) is no longer the crucial feature that allows us to distinguish those beliefs that have to be accounted for in terms of TA from those that have to be accounted for in terms of ALA, then it seems that the two accounts become almost interchangeable. For example, we could rewrite the TA examples of fire/predator in ALA-terms without losing anything in terms of the explanation of the belief-forming process in respect of predators or fires:

> TA*. First consider a creature's manifest surroundings – for example, that there is a fire raging in front of it, or a predator rushing toward it. Some belief-forming mechanisms serve to link certain kinds of circumstances – the presence of a fire/predator – with adaptive responses – the immediate escape behavior.

One would be hard-pressed to deny that TA* is equivalent to the ALA example of the hot surface. TA* does not make reference to any truth that has to be grasped in order to be able to survive, except in the minimal sense of detecting a peculiar state of affairs within the environment. But it would not be fair to argue that this ability in detecting the actual presence of a fire/predator is able to account for the distinction between TA and ALA, since even in the ALA example given by Street herself, as we have seen, it is relevant to detect the actual presence of the hot surface. If even ALA-beliefs can reliably track the state of affairs of the environment and be, in such a minimal sense, truth-sensitive, then it is not easy to understand Street's claim about the need for an organism to have TA-beliefs in order "to grasp" relevant truths about the environment and then survive.

Therefore, Street's claim that ALA applies to the moral domain − since in the moral domain it is not relevant to track truths − while TA applies to other domains is not supported by the examples she gives in order to describe ALA and TA. Such examples are not neatly distinguishable, since a truth-tracking ability also figures in ALA examples. In

this sense, the debunker is epistemically incoherent: she seems not to be able to coherently assign TA or ALA to different sets of beliefs.

The same problem occurs in another interesting debunking position, namely that developed by Richard Joyce. According to Joyce (2006), we have no reason to suppose that our moral beliefs are true insofar as the evolutionary explanation of our moral beliefs need not presuppose that there was any moral fact to grasp in the ancestral environment in which those beliefs evolved. Joyce adopts a criterion to discriminate between TA and ALA according to which we should be realist about the beliefs of a domain D if and only if those beliefs figure into the best explanation of such domain. This is ultimately the same criterion adopted by Street ("one will need to posit in one's best explanation that there were *indeed* fires, predators, and cliffs"). Since moral truths do not figure in our best explanations of the moral domain, we should not be realist about moral beliefs.[11] So, Joyce adopts a sort of ALA to account for moral beliefs.

The problem is that Joyce arrives at a very different conclusion with regard to mathematics:

TA2.   False mathematical beliefs just aren't going to be very useful. Suppose you are being chased by three lions, you observe two quit the chase, and you conclude that it is now safe to slow down. The truth of "1 + 1 = 2" is a background assumption to any reasonable hypothesis of how this belief might have [evolved]. (Joyce 2006, 182).

Thus, Joyce appeals to TA to account for mathematical beliefs. Hence, he supports mathematical realism, the view according to which mathematical abstract objects are causally inert mind-independent entities (Balaguer 2009). This is a crucial point, since there is a long tradition of comparing moral knowledge to mathematical knowledge, and many authors think that the comparison shows that one should accept (or refute) moral realism and mathematical realism for the same reasons (Clarke-Doane 2014). For example, Hilary Putnam states that arguments "for 'antirealism' in ethics are virtually identical with arguments for antirealism in the philosophy of mathematics" (Putnam 2004, 1). The core idea is that if one accepts mathematical realism and denies moral realism, one risks being incoherent. Things do not change if one takes a Darwinian perspective on morality. As Justin Clarke-Doane points out, the evolutionary challenge for "moral realism is equally a challenge for mathematical realism" (Clarke-Doane 2012, 314).[12]

Let's now turn to Joyce's TA2. Despite what Joyce claims, if we compare the aforementioned case of the hot surface (i.e. Street's example of ALA) and Joyce's case of

---

[11]   We will not analyze here the reasons Joyce offers to justify his claim that moral truths do not figure in our best explanations of the moral domain.

[12]   On this issue, see below, sections 3.3 and 4.

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

the three lions, it is not clear at all why in the former case we should refer to ALA in order to explain our capacity of performing an adaptive response, while in the latter we should refer to TA in order to explain our capacity in coping with the world. In other words, it is not clear why in the former case it is not necessary to presuppose the existence of the content of the belief to explain our behavior, while in the latter case it is necessary. If our ability to immediately withdraw our hand is not due to the actual existence of a hot surface, it is not clear why our ability in continuing to escape should instead be dependent on the existence of the numbers "1" and "2", and on the fact that "1 + 1 = 2" has to be regarded as a truth in a realist sense. After all, even if we accept that "1 + 1 = 2" is an evolved belief, because it is useful in coping with the world, this doesn't mean that "1" and "2" are mind-independent entities as Joyce's TA2 seems to imply. Indeed, we can rewrite the three lions example as follows, without losing anything in terms of the explanation of the formation of the adaptive response:

> TA2*. Consider a creature's manifest surroundings – for example, that there are several lions that are chasing it. After a while, some lions give up the chase. Some belief-forming mechanisms serve to link certain kinds of circumstances – i.e. there could still be some lions involved in the chase – with adaptive responses – i.e. continue the escape behavior.

It is difficult to deny that TA2* is equivalent to ALA. But this leads to the question: Is the figuring of numbers in our explanation of some creature's behavior a reliable criterion for determining whether such creature's beliefs about numbers have to be accounted for in ALA-terms or in TA-terms?

The problem with Joyce's perspective is precisely that he considers it to be a reliable criterion for discriminating between beliefs that have to be accounted for in TA-terms and those that have to be accounted for in ALA-terms, whether those beliefs figure in our best explanations. Details are not relevant here, but Clarke-Doane has clearly pointed out that, especially in the case of mathematical beliefs, the presence of those beliefs in our best explanations and our capacity to provide a plausible explanation for such presence come apart. Generalizing, "*we may be able to show that the contents of our beliefs of a kind, D, figure into their best explanation despite having no idea how those beliefs manage to reliably align with the D-truths*" (Clarke-Doane 2015, 99).

But if we cannot give a plausible account of *how* the D-truths and our D-beliefs are related, we can justify neither (1) the claim that those beliefs really correspond to those independent truths in a non-accidental way, nor (2) the claim that our view can be regarded as naturalist. In other words, if we do not know how D-beliefs are formed, we cannot know whether or not the following counterfactual, which describes the correlation that must

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

obtain between the D-beliefs and D-truths in order for us to take a realist attitude toward a given domain, holds: had the truths of domain D been different, our D-beliefs would have been correspondingly different (Clarke-Doane 2015, 88).

To recapitulate, according to Joyce, in order to decide whether we have to deal with TA to account for a domain D, we should be able to show whether or not D-beliefs could have easily been different. We cannot further develop this point here, but since, at least in the case of mathematics, we have "*no idea how those beliefs manage to reliably align with the D-truths,*" it is not easy to determine whether or not mathematical beliefs could have been different. Indeed, this is a very controversial issue. For example, Nagel states that whatever else we may be able to imagine as different, "including the possibility that we ourselves should be incapable of thinking that 2 + 2 = 4, none of it tends to confer the slightest glimmer of possibility on that proposition's failing to be true" (Nagel 1997, 55), because we "cannot conceive of 2 + 2 being equal to 5" (Nagel 1997, 63). Contra Nagel, Dehaene states that "at our scale, the world is mostly made up of separable objects that combine into sets according to the familiar equation 1 + 1 = 2. This is why evolution has anchored this rule in our genes", but "perhaps our arithmetic would have been radically different if, like cherubs, we had evolved in the heavens where one cloud plus another cloud was still one cloud" (Dehaene 1997, 249). So, it seems that, according to Dehaene, not only can we conceive of 2 + 2 being equal to 5, but also, if the human species had evolved in a different world, "conceiving of 2 + 2 being equal to 5 would have been most natural. In that case, presumably, Nagel would have said that we could not conceive of 2 + 2 being equal to 4" (Cellucci 2017, 122).

We will not enter into this debate here. However, given the high degree of controversy that often results from the consideration of whether or not the beliefs of a certain domain could have been different, merely arguing that the contents of our beliefs figure into their best explanation does "nothing evidently to show that they could not have easily been different" (Clarke-Doane 2015, 95). So, even if the beliefs of a certain domain figure in our best explanations, this does not suffice to decide whether we have to adopt TA or ALA in order to account for those beliefs. This is the reason why Joyce risks being epistemically incoherent in supporting ALA for moral beliefs and TA for mathematics.

Both the truth-tracking requirement and the requirement of figuring in our best explanations seem inadequate as criteria for assigning ALA or TA to a certain domain, since they are unable to neatly discriminate between ALA-beliefs and TA-beliefs, at least in light of the examples proposed by Street and Joyce themselves. As noted above, if the debunker is unable to provide a principled way of discriminating ALA-beliefs from TA-beliefs, and she maintains that domain D1 has to be accounted for in terms of ALA, while some other domain D2 has to be accounted for in terms of TA, she risks being epistemically incoherent in her choice.

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

3.3. Inaccessible Domains and Naturalism

In order to handle the epistemic incoherence objection, we consider two crucial and related aspects which may threaten the debunking position: the relation that is assumed to obtain between natural selection and truth, and the existence of a problem of accessibility for the beliefs we form in some domains.[13]

The epistemic incoherence objection may be summarized as follows: the debunker either (1) claims that natural selection is an aimed-at-truth process which is able to epistemically justify our beliefs, or (2) she denies such claim. The debunker also needs to specify whether (3) the domains are all equal with respect to (1) and (2), or (4) they are not equal.

As we have seen before, generally the debunker maintains (1) and (4), and then she has to face the epistemic incoherence objection: since there is no principled way of neatly distinguishing between domains whose beliefs can be epistemically justified by natural selection, and domains whose beliefs cannot be epistemically justified by natural selection, she risks being incoherent in claiming that some domains can be justified by relying on natural selection, while some other domains cannot be so justified. If the debunker wishes to avoid being incoherent, she has two options: either (a) she assumes TA to be the appropriate way to account for every domain, i.e. she maintains (1) and (3), or (b) she assumes ALA to be the appropriate way to account for every domain, i.e. she maintains (2) and (3).

If the debunker assumes (a) – i.e. she adopts a realist attitude for every domain – she risks being compelled to adopt a realist stance also for *a priori* domains, i.e. domains which are afflicted by a problem of accessibility, whose realistically interpreted beliefs the debunkers usually try to debunk (think of morality, for example), precisely because those beliefs cannot be accounted for in naturalist terms. Thus, if the debunker adopts a realist attitude toward *a priori* domains, she risks being unable to claim that her position is a naturalist position. If, on the contrary, the debunker takes (b) – i.e. she adopts a 'deflationary' attitude toward every domain – then she is certainly able to claim that her position is a naturalist one. This comes at a cost, however, since the debunker is compelled to give up a realist attitude toward the beliefs of some domains that are usually realistically

---

[13] The access problem, first raised in the philosophy of mathematics by Paul Benacerraf (1973), is now thought to arise in many other domains. It is the problem of justifying the claim that our D-beliefs align with the D-truths of a given domain D, if D is regarded as an *a priori* domain, i.e. a domain whose objects cannot *in principle* be empirically investigated. See Clarke-Doane 2016; Jonas 2016. For more on this issue, see section 4.

interpreted. This may involve the adoption of some revisionary views of how we conceive of those beliefs.

Indeed, there is a sort of trade-off between taking a realist attitude toward *a priori* domains, and as such not having to be revisionary on some central epistemological issues, and taking naturalism to be our most fundamental meta-philosophical commitment. The debunker has to choose between realism for *a priori* domains and naturalism. In our view, given the role that the commitment to naturalism plays in the elaboration of many EDAs, and the role that Darwinism plays in the development of a naturalist perspective, the more promising route for the debunker is to adopt a deflationary stance for every domain, thus avoiding the epistemic incoherence objection, and maintaining naturalism.

According to some authors, this option may lead to skepticism. In the next section, we will set out our proposal in greater detail. We will argue that it does not necessarily lead to skepticism, and that, as such, it is a viable route for the debunker to take.

4. Debunkers and Scepticism

The debunker, we argued, may avoid the epistemic incoherence objection if she adopts ALA to account for the evolutionary root of the belief-forming process for every set of beliefs. This means that the debunker should take natural selection as unable to justify the claim that our D-beliefs align with D-truths for every domain D. In this view, the formation of D-beliefs may be related to the adaptive advantage that the possession of these beliefs afforded our ancestors, but it cannot guarantee that the content of D-beliefs is made true by some independent fact, or truth, or object.

It may be objected that this move would lead to the skeptical result that it is impossible to verify that the content of any belief that we produce is true. But it just means that taking the evolutionary route for providing an epistemic justification to our beliefs is problematic, if one wants both (1) to take a naturalist stance, and (2) to avoid the epistemic incoherence objection.

We think that in order to better spell out our proposal, it is important to distinguish (a1) the fact that a subject S has some D-beliefs about a domain D, and (a2) the ability in dealing with the world that those D-beliefs confer to S, from (b1) the beliefs that S has about the nature of such D-beliefs, and (b2) the beliefs that S has about the reasons why those D-beliefs give her such an ability in dealing with the world.

In this regard, consider sight, i.e. the ability to see. Sight gives us an ability in dealing with the world (a2), and allows us to form beliefs related to what we see and about the possibility of seeing that may well be regarded as reliable to some extent (a1). Nevertheless, for many millennia humans have known very little about how sight was

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

possible (b1), and many ideas we humans have proposed to explain this phenomenon proved untenable in the light of successive scientific inquiry (b2).

We may conceive of our ability to see as evolutionarily rooted. Thus, natural selection may well have equipped us with the ability to form reliable (at least to some extent) beliefs through our innate ability to see, and these beliefs proved very useful in dealing with the world. Nevertheless, we would not draw the conclusion that our innate ability to see allowed us to form reliable beliefs about how it is that we can see, or about how human sight works. On the contrary, it is starting from our current scientific knowledge that we can assess whether (and to what extent) our innate ability to see allows us to form reliable beliefs.

In a similar vein, consider mathematics. Mathematicians are supposed to have highly justified beliefs about mathematics (a1), and mathematics certainly helps us in dealing with the world (a2). But this does not mean that the beliefs that philosophers of mathematics or mathematicians have about what mathematics is (b1), or about the reasons why mathematics proves so helpful in dealing with the world (b2), are justified or reliable.

In other words, even if natural selection gave us the ability to produce some reliable D-beliefs in some domain D, because these beliefs were useful in dealing with our ancestors' environment, this does not mean that any belief we may produce about D or about how D-beliefs are produced is reliable or justified.

It is important to clarify that this view does not amount to advocating some sort of evolutionary skepticism; rather, it aims at developing a coherent debunking naturalist stance. Naturalism cannot be regarded as necessarily implying skepticism, unless one assumes an old-fashioned, armchair epistemological perspective, which is intimately anti-naturalist. Vlerick and Broadbent (2015) clearly describe what is at stake here. We are not denying the reliability of any belief we form on evolutionary grounds. We are claiming that we have to be cautious in thinking that natural selection can provide an adequate epistemic justification for any belief we form, especially in domains so remote from our ordinary experience, including those beliefs that constitute the greatest part of our scientific knowledge. Indeed, the "main task awaiting the naturalist is not so much debunking the evolutionary sceptic – who denies us justification of even our most basic […] beliefs on the grounds that they are the product of faculties shaped by a natural evolutionary process – but to provide our best scientific theories with epistemic justification" (Vlerick and Broadbent 2015, 190).[14]

---

[14] On this strategy aimed at 'debunking' the evolutionary skeptic, see e.g. Wilkins and Griffiths (2013), who state that to "defeat evolutionary skepticism, true belief must be linked to evolutionary success in such a way that selection will favour organisms which have true beliefs" (Wilkins and Griffiths 2013, 134).

According to Vlerick and Broadbent, often "it is tacitly assumed that rebutting the sceptic by establishing that natural selection shaped our cognitive faculties" (Ibid.), which are able to give us reliable (at least to some extent) beliefs in certain (ecologically relevant) contexts, would amount to providing our best scientific theories with epistemic justification. This "tempting, but fallacious, line of thought" (Ibid.) states that since our scientific beliefs are the product of evolved faculties, and given that these faculties provide us with reliable beliefs in ecologically relevant contexts, they can be expected to afford us reliable beliefs in general. Vlerick and Broadbent describe this argument as follows:

(1) Our cognitive faculties are reliable.

(2) Scientific beliefs are produced by our cognitive faculties.

(3) Scientific beliefs are reliable.

As they point out, the problem with such an argument is the following: What could warrant (1), other than scientific evidence for the reliability of our cognitive faculties? But if we accept scientific evidence for the reliability of our cognitive faculties, we implicitly accept (3), the conclusion. If we reject (3), we will reject the scientific evidence for (1). Thus, "the argument is epistemologically mute" (Ibid., 191). As we said, it is starting from our current scientific knowledge that we can assess whether our innate ability to produce some kind of beliefs allows us to form reliable beliefs, and not the other way around.

Returning to the proposal of adopting ALA in order to account for the beliefs-forming process in every domain, this does not rule out that our evolved faculties may give us some true beliefs. If we intend 'truth' in some thin metaphysical sense, as in the minimal sense according to which truth is understood with regard to our ability of pre-theoretically dealing with the world, then ALA is able to account for the beliefs (and the behaviors) involved in such a pragmatic and minimal construal of 'truth', as demonstrated in the case of the hot surface analyzed above.

If instead 'truth' is intended in some thick metaphysical sense, e.g. as a correspondence between the content of our beliefs and the world itself, then it will be unfair to accuse the debunker of being unable to secure our knowledge and being prone to skepticism. Indeed, in order to fairly advance such an accusation, the 'anti-debunker' must be able to demonstrate that it is possible to both assert such a correspondence and maintain a naturalist stance. But to date, providing such a demonstration has proven difficult.[15]

---

[15] Consider, for instance, the idea that truth is correspondence. In order to give a rigorous construal of the concept of truth as correspondence, scientific realists usually conceive of truth in Tarskian terms (Ruttkamp 2002). The correspondence relation can be made precise in Tarskian terms because, following Tarski, it is possible to show that an isomorphism obtains between two

Now, it may be objected that if the D-beliefs we produce prove useful in dealing with the world, this means that they are true. Thus, even if we do not have any plausible account of how it is that our D-beliefs align with D-truths, we nevertheless may safely regard those beliefs as true. So, asking for a plausible account of how we form those beliefs would be unnecessary. Contrary to this objection, the fact that some beliefs prove useful in dealing with the world provides insufficient support for the conclusion that those beliefs are true. First of all, even false beliefs may prove useful in dealing with the world. Think of Newtonian Gravitation. We now regard it as false, strictly speaking, but nevertheless it proved (and remains to this day) very useful in dealing with the world (Bueno 2017). Thus, even if a belief is useful in dealing with the world, this does not guarantee that such a belief is true. Secondly, if we do not have a plausible account of how we form some D-beliefs, even if those beliefs prove useful in dealing with the world, this does not suffice to justify a realist attitude toward D-beliefs, especially when taking a realist attitude amounts to claiming that the content of such D-beliefs is determined by an independent realm of D-facts or D-truths causally disconnected from our world (see below). Think again of mathematics. Mathematics is certainly useful in dealing with the world. And there is strong agreement on the reliability of mathematical results among mathematicians. That this fact should be interpreted as indicating that the content of mathematical beliefs is determined by a realm of mathematical facts or truths causally disconnected from our world is instead a very controversial metaphysical issue. Moreover, even if we could determine with certainty that there is not a mathematical realm causally disconnected from our world, we would continue to claim that mathematics proved very useful in dealing with the world, because it in fact proved useful.

It is worth noting that our proposal is aimed at avoiding the situation in which a naturalist debunker may be compelled to adopt a realist stance toward inaccessible domains, i.e. domains that have to face the access problem (Jonas 2016). Taking a realist stance toward those domains, indeed, would be incompatible with a naturalist stance.

---

mathematical structures, and the notion of isomorphism is considered to rigorously instantiate the notion of correspondence. But if we try to adopt this line of reasoning, we should admit that the world is a mathematical structure. Indeed, an isomorphism can only obtain between two mathematical structures (Halvorson 2012), and if an isomorphism is claimed to obtain between mathematics and the world, then the world has to be regarded as a mathematical structure. For example, Tegmark states that from "the definition of a mathematical structure […], it follows that if there is an isomorphism between a mathematical structure and another structure […], then they are one and the same. If our external physical reality is isomorphic to a mathematical structure, it therefore fits the definition of being a mathematical structure" (Tegmark 2008, 107). This is obviously a sort of Pythagoreanism, which is certainly a legitimate philosophical position, but it is normally considered to be at odds with a naturalist stance (Steiner 1998).

This may be clearly demonstrated by considering the consequences of not adopting the strategy for avoiding the epistemic incoherence objection we are advocating. Indeed, since there is not a principled way to neatly distinguish between domains whose beliefs can be epistemically justified by natural selection and domains whose beliefs cannot be epistemically justified by natural selection, unless the debunker adopts ALA for every domain, she risks being unable to discriminate between domains which are afflicted by the access problem and those which are not afflicted by this problem. If anti-naturalists provide some EAs for inaccessible domains, the debunker may not be able to resist that argument.

Consider morality. Moral realism is usually regarded as an anti-naturalist perspective. The problem of moral realism is a problem of accessibility. In this respect, morality is analogous to other domains that have traditionally been considered as afflicted by access problems, e.g. modality, mathematics, and logic (see, e.g. Jonas 2016; Clarke-Doane 2016). Thus, the main challenge for moral realists is an epistemological challenge, which is analogous to the famous Benacerraf-Field challenge for mathematical realists (Clarke-Doane 2014; 2105).[16]

If D is a domain of *a priori* truths, the access problem for realism "about D consists in the apparent impossibility of explaining how knowledge of the D-truths is possible, if these truths are understood realistically" (Jonas 2016, 1). In this context, realism is understood "as committed to the claim that the content of moral, mathematical, […] and logical beliefs describes properties and facts which are isolated from causally efficacious properties" (Woods 2016, 3).

According to Clarke-Doane (2012; 2014; 2015), for a domain D whose *a priori* truths are interpreted realistically, the access problem undermines belief in the D-truths to the

---

[16] Cf. Benacerraf 1973, 674: "[O]n a realist […] account of mathematical truth our explanation of how we know the basic postulates must be suitably connected with how we interpret the referential apparatus of the theory. […]. [But] what is missing is precisely […] an account of the link between our cognitive faculties and the objects known. […]. We accept as knowledge only those beliefs which we can appropriately relate to our cognitive faculties." See Clarke-Doane (2016). Benacerraf's epistemological challenge to mathematical Platonism has been criticized because it assumes the causal theory of knowledge, which nowadays is discredited among epistemologists. But Benacerraf's argument may be raised against Platonism without assuming the causal theory of knowledge, as Field maintains (Field 1989). On this issue cf. Baron 2015, 152: "Field's version of the access problem focuses on mathematicians' mathematical beliefs. The mathematical propositions that mathematicians believe tend to be true. If platonism is correct, however, then these propositions are about mathematical objects. So, the mathematical beliefs held by mathematicians […] are reliably correlated with facts about such objects. The challenge facing the platonist, then, is to provide an account of this reliable correlation."

extent that it gives us reason to doubt the safety or sensitivity of our D-beliefs. Safety and Sensitivity conditions are usually characterized as follows:

Sensitivity: If *p* were false, S would not believe that *p*.

Safety: In all nearby worlds where S believes that *p*, *p* is not false.

Clarke-Doane develops an argument according to which, if (1) D-truths are necessary truths, and (2) our D-beliefs are evolutionary determined, i.e. they are instilled in us by natural selection, then our D-beliefs may well be regarded as sensitive and safe. As to (1), if the D-truths are necessarily true, then belief in the D-truths is trivially sensitive. As to (2), if D-beliefs have been evolutionarily determined, they could not have easily been wrong, thus belief in the D-truths is safe. Given that evolutionary considerations support the claim that D-beliefs are sensitive and safe, in Clarke-Doane's view, this strategy may allow us to deflect the access problem for *a priori* domains. If this argument holds, the debunker may find it difficult to defend her naturalist attitude. Indeed, she would find herself committed to some form of modal realism, which is usually refuted by naturalists.

Let us try to explain this point by unpacking Clarke-Doane's argument a bit. As to (2), it rests on the standard way in which EAs are supposed to provide epistemic justification to our innate beliefs, i.e. if our innate beliefs were easily wrong, we would not have survived. We survived, so our innate beliefs could not easily have been wrong. We will not question this kind of argument here, at least for argument's sake. Instead, we will focus on (1), i.e. if the D-truths are necessarily true, then belief in the D-truths is trivially sensitive. The main problem with this line of reasoning is that (1) follows from "the fact that the first part of sensitivity conditional – S's belief that *p* is sensitive iff, if *p* were false, S would not believe that *p* – is never actualized for necessary truths on a standard semantics" (Jonas 2016, 6). In other words, if D-truths are necessary truths, then it is impossible to believe they are not true. But how can we affirm that D-truths are necessary truths? Usually, the concept of necessary truth is spelled out in terms of possible worlds. Necessary truths are true in all possible worlds. How can we justify this way of conceiving of necessary truths? 'Necessary truths are true in all possible worlds' is a modal claim, and modality is a domain usually regarded as an *a priori* domain, i.e. a domain afflicted by the access problem. Now, how can we justify the claim that we know that a given domain D is a domain of necessary truths?[17] If we know that D-truths are necessary truths in virtue of the truths of another

---

[17]   The analogy with mathematics is instructive even with respect to modality. Indeed, mathematics is seen to provide support for anti-naturalism. If mathematics provides an instance of genuine *a priori* knowledge, and we do not possess any naturalist account of mathematics, this means that other domains may well be regarded as instances of *a priori* knowledge. But this view rests on the

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

domain, say the modal domain M, then we have to assess whether our M-beliefs are true or not. If we take an antirealist attitude toward modality, then our M-beliefs do not necessarily correspond to some independent fact or truth. Thus, our M-beliefs about D-truths may be false. This would lead us to the quite paradoxical situation in which our knowledge of a necessary D-truth is dependent on our having a fallible M-belief. If, on the contrary, we take a realist attitude toward modality, we fall back into the problem of justifying our access to modal truths. Indeed, modal realists view possible worlds as causally isolated from ours. In this case, we would find ourselves in the uncomfortable position whereby our knowledge of an inaccessible D-truth is dependent on an inaccessible M-truth. At this point, Clarke-Doane can reiterate an EA for modal truths analogous to the one he advocated for in the case of moral truths. If the debunker accepts EAs in some domains (i.e. if she adopts TA for some domains, since, as shown above, EAs are equivalent to TA), and there is no principled way of discriminating between TA-domains and ALA-domains, then she will be unable to refuse Clarke-Doane's EA for modality. Thus, she risks finding herself being committed to some anti-naturalist view. In fact, in the context of modal epistemology, adopting a realist attitude toward modal truths usually amounts to adopting some formulation of possible worlds modal realism, a view that is usually deemed incompatible with a naturalist stance (Vaidya 2016; Bueno 2017).[18]

---

simplistic idea that mathematical statements may *really* be proved to be true. In fact, things are more complicated. According to Bueno and Shalkowski (2004), for instance, as in mathematics, due to Gödel's results, we are generally unable to prove with certainty that the axioms of the theory we are dealing with are true, and thus that the theorems that we derive from such axioms are actually true, when dealing with modality our modal knowledge may be of the same kind, i.e. knowledge whose truth depends on whether the metaphysical assumptions from which we start are true, but we are unable to prove whether such assumptions are actually true. Cf. Bueno and Shalkowski 2004, 97-98: "If the analogy with mathematics is taken seriously, it may actually provide a reason *to doubt* that we have any knowledge of modality. One of the main challenges for platonism about mathematics comes from the epistemological front, given that we have no access to mathematical entities – and so it's difficult to explain the reliability of our mathematical beliefs. […]. Reasons to be skeptical about a priori knowledge regarding mathematics can be easily 'transferred' to the modal case, in the sense that difficulties we may have to establish a given mathematical statement may have a counterpart in establishing certain modal claims. For example, how can we know that a mathematical theory, say ZFC, *is consistent?* Well, we can't know that in general; we have, at best, relative consistency proofs. And the consistency of the set theories in which such proofs are carried out is far more controversial than the consistency of ZCF itself, given that such theories need to postulate the existence of inaccessible cardinals and other objects of this sort." On related issues, see Cellucci (2017).

[18]   The adoption of possible worlds modal realism amounts to assuming that there is something "like a realm of metaphysical possibility and necessity that outstrips the possibility and necessity that

In this respect, it may be objected that some authors claim that we can distinguish among *a priori* domains. Woods (2016), for instance, claims that morality is more dispensable than mathematics and logic, because, even if for all those three domains we lack a plausible naturalist account, morality and moral beliefs may play no role in the process of selection of our best explanations, while mathematical and logical beliefs are indispensable for the very definition of what constitutes a 'best explanation'. Thus, those domains would be in some way immune to certain debunking maneuvers. This view would allow the debunker not to be revisionary with respect to domains that are usually realistically interpreted.

The problem with this argument is that it rests on the assumption that the content of our mathematical and logical beliefs, in order to allow us to select what is the best explanation in a certain context, have to be regarded as made true by some inaccessible facts or truths. But this assumption is unjustified. As noted, even if we could demonstrate that there is no inaccessible realm of mathematical facts, we would continue to maintain that mathematics is useful in dealing with the world, because mathematics proved in fact useful in dealing with the world. Whether or not the content of our mathematical beliefs is made true by some inaccessible realm is irrelevant for assessing whether mathematics proved useful or not. According to Woods, we cannot construct a debunking argument for mathematics and logic, because mathematics and logic are necessary in order to assess whether such debunking explanation is the best explanation for our possession of mathematical and logical beliefs: "the notion of best in best explanation is not independent of our mathematical and logical beliefs – our conclusions about which explanation is best will depend on what mathematical and logical theory we accept in the background" (Woods 2016, 13). But this argument assumes that mathematical and logical beliefs must be true in order to lead us to draw conclusions about which explanation is best. But just as in the past some theory we nowadays regard as false proved useful in dealing with the world, so it could be the case that mathematical and logical beliefs prove useful in helping us draw conclusions about which explanation is best independently of whether they are true. Thus, the fact that some beliefs may be constitutive of some inferential process we adopt in dealing with the world is insufficient to claim that we need not also confront the access problem for those beliefs. The inaccessible domains are all equal from a naturalist perspective.

5. *Conclusion*

science deals with, but this is exactly what naturalists should not be willing to concede" (Morganti 2016, 87).

Focusing on the cutting-edge debate on EDAs in metaethics, we have discussed an epistemological issue, i.e. the epistemic incoherence objection, which affects those arguments and whose interest extends far beyond metaethics. The main thrust of our argument is that if one argues, as the debunker generally does, that Darwinism has a corrosive impact on metaethics, such a corrosive impact cannot be limited to metaethics. We have shown that the effort to neutralize Darwinism in domains besides morality risks not only being epistemically incoherent, but also embracing forms of non-naturalism. Finally, we have also argued that a coherent debunker position does not necessarily imply skepticism. In different philosophical domains one can find several remarkable positions, which can easily be recognized as compatible with the perspective outlined here. Think of the moral domain, for instance. Several philosophical perspectives are available (e.g. Gibbard 1990; Severini 2016), which are not committed to the existence of some independent moral world populated by moral truths. Consider also mathematics. Nominalism is a respectable philosophical view, which contrasts with mathematical Platonism precisely because it denies that we have to commit ourselves to the existence of inaccessible mathematical objects. In turning to the philosophy of science, we may find, e.g. scientific realism confronting epistemic instrumentalism (Stanford 2016a; 2016b). Epistemic instrumentalism contends that we should refrain from claiming that scientific theories tell us the truth about inaccessible reality, and we should be content in regarding those theories as useful and powerful tools to cope with the world. These positions provide examples of how the debunker can well account in ALA-terms for the evolutionary roots of the belief-forming processes in all those domains, and this would lead neither to skepticism, nor to any explanatory loss. Indeed, *if* those positions are legitimate non-skeptical positions, and they do not adopt a realist attitude toward the beliefs we form in their domains, then the debunker who adopts ALA in order to account for the evolutionary roots of those beliefs is not leaving out anything relevant in her explanation.

The key question is: Are all those positions which avoid taking a realist stance in those domains prone to skepticism? We think that the answer is in the negative, unless it is assumed that one can only defeat skepticism if one adopts a realist stance toward every domain (see, e.g. Parfit 2011). If one takes a naturalist stance, indeed, one will be unwilling to adopt a realist attitude (at least) toward some domains, especially the aforementioned inaccessible domains. Thus, the idea that one may defeat skepticism only if one adopts a realist stance toward every domain is not available to the naturalist. Obviously, this worry only affects those who are committed to naturalism in the first place. Nevertheless, since Darwinism is crucial for naturalism, and the supporters of EDAs are usually committed to naturalism, the idea that only if one adopts a realist stance toward every domain can one defeat skepticism is also unavailable to any Darwinian debunker.

We have tried to delineate a viable route for the debunker. Developing a fully coherent debunker position in any detail is beyond the scope of this paper. Nevertheless, revising some of our old-fashioned philosophical views is one of the most fruitful insights of Darwinism. As such, we suggest that if the universal acid cannot be contained, not try to contain it and let it flow.

References

Artiga, M. (2015). Rescuing Tracking Theories of Morality. *Philosophical Studies*, 172 (12): 3357-3374.

Balaguer, M. (2009). Realism and anti-Realism in Mathematics. In: D. Gabbay, P. Thagard, & J. Woods (Eds.), *Philosophy of Mathematics*. Amsterdam: Elsevier: 117-151.

Baron, S. (2015). Mathematical Explanation and Epistemology: Please Mind the Gap. *Ratio*, 29(2): 149-167.

Benacerraf P. (1973). Mathematical Truth. *The Journal of Philosophy*, 70(19): 661-679.

Bogardus, T. (2016). Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument. *Ethics*, 126(3): 636-661.

Boudry, M., & Vlerick, M. (2014). Natural Selection Does Care about Truth. *International Studies in the Philosophy of Science*, 28(1): 65-77.

Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Bueno, O. (2017). The Epistemology of Modality and the Epistemology of Mathematics. In: B. Fischer, & F. Leon (Eds.), *Modal Epistemology After Rationalism*. Cham: Springer: 67-83.

Bueno, O, & Shalkowski, S. (2004). Modal Realism and Modal Epistemology: A Huge Gap. In: E. Weber, & T. De Mey (Eds.), *Modal Epistemology*. Brussels: Koninklijke Vlaamse Academie van Belgie: 93-106.

Cellucci, C. (2017). *Rethinking Knowledge: The Heuristic View*. Dordrecht: Springer.

Clark, K.J. (2016). Naturalism and its Discontents. In: K.J. Clark (Ed.), *The Blackwell Companion to Naturalism*. Oxford: Blackwell: 1-15.

Clarke-Doane, J. (2012). Morality and Mathematics: The Evolutionary Challenge. *Ethics*, 122(2): 313-340.

Clarke-Doane, J. (2014). Moral Epistemology: The Mathematics Analogy. *Noûs*, 48(2): 238-255.

Clarke-Doane, J. (2015). Justification and Explanation in Mathematics and Morality. In: R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics. Volume 10*. Oxford: Oxford University Press: 80-103.

Clarke-Doane, J. (2016). What Is the Benacerraf Problem? In: Pataut, F. (Ed.), *Truth, Objects, Infinity. New Perspectives on the Philosophy of Paul Benacerraf*. Cham: Springer: 17-43.

De Cruz, H., Boudry, M., De Smedt, J., & Blancke, S. (2011). Evolutionary Approaches to Epistemic Justification. *Dialectica*, 65(4): 517-535.

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

History and Philosophy of the Life Sciences, DOI 10.1007/s40656-017-0154-1.

Dennett, D. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

Enoch, D. (2010). The Epistemological Challenge to Metanormative Realism: How Best to Understand it, and How to Cope with it. *Philosophical Studies*, 148(3): 413-438.

Field, H. (1989). *Realism, Mathematics and Modality*. Oxford: Blackwell.

Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge (MA): Harvard University Press.

Graber, A. (2012). Medusa's Gaze Reflected: A Darwinian Dilemma for Anti-Realist Theories of Value. *Ethical Theory and Moral Practice*, 15(5): 589-601.

Halvorson, H. (2012). What Scientific Theories Could Not Be. *Philosophy of Science*, 79(2): 183-206.

Hodge, J., & Radick, G. (2003). Introduction. In: J. Hodge, & G. Radick (Eds). *The Cambridge Companion to Darwin*. Cambridge: Cambridge University Press: 1-14.

Jonas, S. (2016). Access Problems and Explanatory Overkill. *Philosophical Studies*, DOI: 10.1007/s11098-016-0807-z.

Joyce, R. (2006). *The Evolution of Morality*. Cambridge (MA): MIT Press.

Kahane, G. (2011). Evolutionary Debunking Arguments. *Noûs*, 45(1): 103-125.

Kitcher, P. (2011). *The Ethical Project*. Cambridge (MA): Harvard University Press.

Kyriacou, C. (2016). Are Evolutionary Debunking Arguments Self-Debunking? *Philosophia*, 44(4): 1351-1366.

Lewens, T. (2005). What is Darwinian Naturalism? *Biology and Philosophy*, 20(4): 901-912.

McKay, R.T., & Dennett, D.C. (2009). The Evolution of Misbelief. *Behavioral and Brain Sciences*, 32(6): 493-510.

Millikan, R. (1984). Naturalist Reflections on Knowledge. *Pacific Philosophical Quarterly*, 65(4): 315-334.

Morganti, M. (2016). Naturalism and Realism in the Philosophy Science. In: K.J. Clark (Ed.), *The Blackwell Companion to Naturalism*. Oxford: Blackwell: 75-90.

Nagel, T. (1997). *The Last Word*. Oxford: Oxford University Press.

Papineau, D., (2016). Naturalism. In: E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, URL: <https://plato.stanford.edu/archives/win2016/entries/naturalism/>.

Parfit, D. (2011). *On What Matters*, Volume 2, Oxford: Oxford University Press.

Putnam, H. (2004). *Ethics Without Ontology*. Cambridge (MA): Harvard University Press.

Railton, P. (1986). Moral Realism. *Philosophical Review*, 95(2): 163-207.

Rosenberg, A. (1996). A Field Guide to Recent Species of Naturalism. *The British Journal for the Philosophy of Science*, 47(1): 1-29.

Rosenberg, A. (2016). Darwinism as Philosophy: Can the Universal Acid Be Contained? In: D. Livingstone Smith (Ed.), *How Biology Shapes Philosophy. New Foundations for Naturalism*. Cambridge: Cambridge University Press: 23-50.

Ruse, M., & Wilson, E.O. (1986). Moral Philosophy as Applied Science. *Philosophy*, 61(236): 173-192.

Ruttkamp-Bloem, E. (2002). *A Model-Theoretic Realist Interpretation of Science*. Dordrecht: Kluwer.

Severini, E. (2016). Evolutionary Debunking Arguments and the Moral Niche. *Philosophia*, 44(3): 865-875.

Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.

Sinnott-Armstrong, W. (1996). Moral Skepticism and Justification. In: W. Sinnott-Armstrong, & M. Timmons (Eds.), *Moral Knowledge? New Readings in Moral Epistemology*. Oxford: Oxford University Press: 3-48.

Sinnott-Armstrong, W. (2015), Moral Skepticism. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), URL = <https://plato.stanford.edu/archives/fall2015/entries/skepticism-moral/>.

Skarsaune, K.O. (2011). Darwin and Moral Realism: Survival of the Iffiest. *Philosophical Studies*, 152(2): 229-243.

Stanford, K. (2016a). Instrumentalism: Global, Local, and Epistemic. In: P. Humphreys (Ed.), *The Oxford Handbook of the Philosophy of Science*. Oxford: Oxford University Press: 318-336.

Stanford, K. (2016b). Naturalism Without Scientism. In: K.J. Clark (Ed.), *The Blackwell Companion to Naturalism*. New York: Wiley-Blackwell: 91-108.

Steiner, M. (1998). *The Applicability of Mathematics as a Philosophical Problem*. Cambridge (MA): Harvard University Press.

Sterpetti, F. (2015). Are Evolutionary Debunking Arguments Really Self-Defeating? *Philosophia*, 43(3): 877-889.

Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies*, 127(1): 109-166.

Stump, D. (1991). Fallibilism, Naturalism and the Traditional Requirements for Knowledge. *Studies in History and Philosophy of Science*, 22(3): 451-469.

Tegmark, M. (2008). The Mathematical Universe. *Foundations of Physics*, 38(2): 101-150.

Vaidya, A. (2016). The Epistemology of Modality. In: E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, URL: <https://plato.stanford.edu/archives/win2016/entries/modality-epistemology/>.

Vavova, K. (2014). Debunking Evolutionary Debunking. In: R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics. Volume 9*. Oxford: Oxford University Press: 76-101.

Vlerick, M., & Broadbent, A. (2015). Evolution and Epistemic Justification. *Dialectica*, 69(2): 185-203.

Wilkins, J.S., & Griffiths, P.E. (2013). Evolutionary Debunking Arguments in Three Domains: Fact, Value, and Religion. In: J. Maclaurin, & G. Dawes (Eds.), *A New Science of Religion*. New York: Routledge: 133-146.

Woods, J. (2016). Mathematics, Morality, and Self-Effacement. *Noûs*, DOI: 10.1111/nous.12157.