1 Introduction

Diego E. Machuca

1.1 Overview

In his *Autobiography*, in the course of describing his religious thinking in the period between October 1836 and January 1839, Charles Darwin writes:

Another source of conviction in the existence of God connected with the reason and not the feelings, impresses me as having much more weight. This follows from the extreme difficulty or rather impossibility of conceiving this immense and wonderful universe, including man with his capability of looking far backwards and far into futurity, as the result of blind chance or necessity. When thus reflecting I feel compelled to look at a First Cause having an intelligent mind in some degree analogous to that of man; and I deserve to be called a Theist.

But then arises the doubt—can the mind of man, which has, as I fully believe, been developed from a mind as low as that possessed by the lowest animal, be trusted when it draws such grand conclusions? May not these be the result of the connection between cause and effect which strikes us as a necessary one, but probably depends merely on inherited experience? Nor must we overlook the probability of the constant inculcation in a belief in God on the minds of children producing so strong and perhaps an inherited effect on their brains not yet fully developed, that it would be as difficult for them to throw off their belief in God, as for a monkey to throw off its instinctive fear and hatred of a snake.

(1958: 92–93)

A similar uncertainty about the reliability of the mind that forms the belief in theism is expressed in a letter to William Graham from July 3, 1881:

Nevertheless you have expressed my inward conviction, though far more vividly and clearly than I could have done, that the Universe is
not the result of chance. 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?

(Darwin 1887: vol. 1, 316)

The reason why the theory of evolution through natural selection provides us with a _prima facie_ defeater for the conclusion of the theistic argument from design is that the mind that constructs such an argument has been developed from that of lower animals, which we do not regard as reliable when it comes to theoretical matters. It could be argued that the belief in the God of theism is to be explained by its being evolutionarily advantageous—just as the monkey’s fear of snakes is to be explained by its being so—rather than by its being the product of a faculty that has been selected for because of its capacity to reach “grand conclusions.”

Although the two quoted passages occur in the context of the discussion of theistic belief, the doubt that Darwin entertains about the reliability of the faculty that produces that kind of belief may spill over into other domains. For insofar as we use the same faculty to draw conclusions or to form beliefs about non-religious matters, the same sort of defeater could in principle arise for moral, logical, or mathematical beliefs. In fact, in the _Descent of Man_, Darwin makes the following remarks about our moral sense:

It may be well first to premise that I do not wish to maintain that any strictly social animal, if its intellectual faculties were to become as active and as highly developed as in man, would acquire exactly the same moral sense as ours. In the same manner as various animals have some sense of beauty, though they admire widely different objects, so they might have a sense of right and wrong, though led by it to follow widely different lines of conduct. If, for instance, to take an extreme case, men were reared under precisely the same conditions as hive-bees, there can hardly be a doubt that our unmarried females would, like the worker-bees, think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters; and no one would think of interfering. Nevertheless the bee, or any other social animal, would in our supposed case gain, as it appears to me, some feeling of right and wrong, or a conscience.

(1871: vol. 1, 73)

Let us suppose that killing one’s siblings or children really is, _ceteris paribus_, morally wrong. If humans had evolved differently because of environmental pressures, we could have believed that such an action is morally right because believing so would have been evolutionarily
advantageous. Alternatively, let us suppose that killing one’s siblings or children really is, *ceteris paribus*, morally right. Perhaps we do not believe it is so because, given the environment in which humans evolved, believing so was evolutionarily disadvantageous. It could then be argued that evolution through natural selection does not care about moral truth but only about reproduction and survival and, hence, that we have a strong reason to call into question the epistemic credentials of our moral beliefs.

The apparent debunking or undermining implications of evolutionary theory of which Darwin himself was at least somewhat aware have received a lot of philosophical attention over the past twenty years. At present, evolutionary debunking arguments (EDAs) are a hot topic in several areas of philosophical research. It is often remarked that an EDA normally consists of two premises, one that refers to the evolutionary origin of certain types of beliefs or certain belief-forming mechanisms and one that makes the epistemic point that the evolutionary forces that shaped those beliefs or mechanisms are not truth tracking. For example, it could be argued that our holding X beliefs can be explained by appealing to the fact that doing so was evolutionarily advantageous for our ancestors and that such an evolutionary advantage did not require our tracking alleged X facts. If so, then, if any X facts do exist, it would be a highly lucky coincidence that our evolutionarily shaped X beliefs matched the X facts. Since we have no epistemic reason to believe that we got lucky, we must suspend judgment about whether our X beliefs are true. It should be observed that, although EDAs typically appeal to biological evolution, they may also appeal to cultural evolution to explain the origin of the targeted beliefs or belief-forming mechanisms.

EDAs are customarily used to undermine the epistemic justification of our X beliefs, but occasionally they are employed to establish that our X beliefs are (probably) false rather than epistemically unjustified inasmuch as the X facts do not exist. Nevertheless, this second use is illegitimate inasmuch as an EDA by itself is unable to establish an ontological conclusion. It can at most be used as a supplement to other arguments that aim to establish a negative ontological conclusion in order to explain, once the conclusion is accepted, the systematic error we make in making X judgments, and hence in holding X beliefs.

The two areas in which EDAs have attracted the most attention are ethics and the philosophy of religion. It is in ethics where in recent years there has been an explosion of interest in EDAs. Among authors advancing moral EDAs, one can distinguish between radical and moderate evolutionary moral debunkers. The former target moral realism of both a naturalist and a non-naturalist stripe (Ruse 1986; Joyce 2001, 2006, 2016; Kitcher 2006; Street 2006; Fraser 2014; Braddock 2016). The two most important radical debunkers are Richard Joyce and Sharon Street. Joyce’s EDA is based on a detailed evolutionary account of morality according to which the formation of beliefs about objective moral
rightness and wrongness may have served to enhance our ancestors’ reproductive fitness independently of whether there existed any moral properties or facts. While Joyce (2001) uses an EDA mainly as a supplement to his arguments for a moral error theory, Joyce (2006, 2016) appeals to an EDA to establish an epistemological skepticism. Joyce (2016) also claims that EDAs place the burden of proof on the moral realist defending the justification of our moral judgments inasmuch as EDAs present a new plausible hypothesis about the origin and epistemic status of those judgments. Street (2006) argues that the fact that the forces of natural selection have had a tremendous indirect influence on the content of our evaluative judgments raises the challenge to explain the relationship between such evolutionary influences and the independent evaluative facts posited by the normative realist. The normative realist then faces a dilemma: either he is forced to embrace a far-fetched moral skepticism, or he must propose a scientifically unacceptable tracking account. Moderate debunkers restrict the debunking scope of evolutionary theory, claiming that evolutionary considerations call into question merely those moral beliefs that are held because everyone accepts them (Parfit 1984), or maintaining that such considerations undermine only non-utilitarian or non-consequentialist normative theories (Greene 2008, de Lazari-Radek & Singer 2012), or constructing an EDA that purports to undermine only naturalistic moral realism (Bogardus 2016).

Regarding the philosophy of religion, EDAs have been discussed to a large extent in connection with the so-called cognitive science of religion (CSR), which explains the origins of religious thoughts, experiences, and practices by appealing to theories, methods, and findings from cognitive, developmental, and evolutionary psychology, as well as from anthropology, archaeology, and the history of religions. Two chief CSR theories have been proposed: whereas some claim that religious belief and behavior were selected for because they enhanced reproduction and survival (e.g., Wilson 2002, Bulbulia 2009, Bering 2011), others contend that religious belief and behavior are by-products (usually called “spandrels”) of traits that were adaptive (e.g., Boyer 2001, Atran 2002, Barrett 2004). Either theory may be appealed to in arguing that religious beliefs are formed and sustained by cognitive mechanisms that do not track religious truths. But it should be noted that not all the proponents of the two theories take them to have debunking implications. It is also worth observing that it has been argued that CSR is actually independent of evolutionary science inasmuch as one can investigate how human cognition informs religious beliefs and practices without explaining why human cognition is the way it is. But CSR and evolutionary science are typically associated because the latter provides a plausible explanation of why our cognitive mechanisms are the way they are (see Barrett 2007).
Although in the literature on evolutionary debunking the main focus of attention has been on EDAs in ethics and philosophy of religion, there are at least three other areas in which discussion of EDAs is of philosophical significance. First, mathematical realism has been called into question on the basis of evolutionary considerations. This has been done particularly by drawing a parallel between mathematics and morality: it has been argued that the evolutionary challenge to moral realism is, despite what most philosophers think, equally a challenge to mathematical realism (Clarke-Doane 2012, 2020). The reason is that one can well imagine that, had the mathematical truths been very different, our mathematical beliefs would have been the same inasmuch as it would still have benefited our ancestors to have the same mathematical beliefs. Second, one could appeal to similar evolutionary considerations to debunk our metaphysical beliefs about ordinary objects. For instance, one could call into question the belief that there are colors on the surfaces of objects or the belief that there really are dogs and apples and not merely atoms arranged dog-wise or apple-wise. Third, although the discussion of the evolutionary debunking of morality, religion, or mathematics falls within the scope of moral, religious, or mathematical epistemology, one may focus on EDAs to explore wider epistemological questions. For example, one may examine whether EDAs can be used to call into question the reliability of our cognitive faculties in general and whether such arguments are self-defeating. One may also look into the general epistemological principles that underlie EDAs in any domain.

Even though EDAs have for several years been a hot topic in philosophy, until now, no collective volume has been entirely devoted to such arguments or has examined EDAs in the five areas of philosophical research mentioned in the foregoing paragraphs. The purpose of the present volume is to fill this gap in the literature. It brings together fourteen original essays that cast fresh light on old problems or propose new lines of inquiry.

1.2 Preview of the Essays

This volume is divided into four parts, dealing with EDAs in ethics, philosophy of religion, philosophy of mathematics, and metaphysics and epistemology. In what follows, I offer a summary of each of the chapters.

The first four essays deal with EDAs in ethics. In “Debunking What?” Hallvard Lillehammer focuses on the practical implications of endorsing the conclusions of EDAs—and genealogical debunking arguments more generally. More precisely, he considers whether one can retain one’s commitment to at least some moral claims despite their alleged lack of truth or justification. According to Lillehammer, although in some cases one cannot debunk the epistemic credentials of certain moral claims without
thereby debunking their moral standing, in others our commitment to certain moral claims is strong enough to withstand the discovery that they are the product of a previously unknown non-rational causal process. The reason is that there are certain things that we morally value no matter how we may actually have come to value them.

In “The Evolutionary Debunking of Quasi-Realism,” Neil Sinclair and James Chamberlain present a novel answer that quasi-realists can provide to a version of the reliability challenge in ethics—which asks for an explanation of why our moral beliefs are generally true—and in so doing, they examine whether evolutionary arguments can debunk quasi-realism. Although reliability challenges differ from EDAs in several respects, there may well be a connection between them. For the explanatory premise of an EDA may state that a particular theory of beliefs of a certain kind does not, or cannot, provide a plausible account of why those beliefs might be generally true, and its epistemic premise may state that, if that is the case, then the beliefs in question have a negative epistemic status or the theory is false inasmuch as, if it were true, it would lead to those beliefs having such a negative epistemic status. The quasi-realist can answer the reliability challenge by claiming that, when we form our moral beliefs through a process of well-informed impartial reflection, we form them in response to the non-moral features of things on which depend the moral features they have. Hence, when we form beliefs by means of such a process, we are most likely forming true moral beliefs.

Street’s EDA against moral realism—and value realism more generally—exploits the idea that it would be a highly unlikely coincidence that our evolutionarily shaped moral beliefs matched the objective moral truths posited by the moral realist, given that such truths seem to be irrelevant in accounting for the selection forces that molded human moral psychology. In “Fine-Tuning the Darwinian Dilemma,” Andreas Mogensen engages with this challenge to moral realism, which he calls “the Coincidence Problem.” He argues for the following two claims. First, pace Street, constructivists are no less committed than realists to acknowledging that only by a coincidence has natural selection favored the evolution of reliably accurate evaluative judgments. Second, if we draw an analogy between the Coincidence Problem and the so-called Fine-Tuning Problem in the philosophy of cosmology, we realize that a commitment to some sort of explanatory coincidence need not be objectionable for the moral realist.

In “Virtue Epistemology and Evolutionary Debunking of Morality,” Michael Klenk maintains that it is possible to formulate a plausible general epistemic principle that may serve as a premise in an EDA against moral beliefs by appealing to recent approaches to knowledge that center on the role of intellectual or epistemic virtues. Relying on the principle he dubs “the achievement conception of undercutting defeat,” Klenk argues that evolutionary explanations of morality can undermine the epistemic
justification of all our moral beliefs if they show that our cognitive success in morality is not sufficiently creditable to the competent use of our cognitive abilities. In his view, not only does the antecedent of that conditional not obtain, but evolutionary explanations actually strengthen the epistemic standing of our moral beliefs by helping us understand how it is that our cognitive abilities enable us to be cognitively successful in morality.

The second part of the volume, devoted primarily to EDAs against religious belief, consists of four essays. As its title indicates, Max Baker-Hytch’s “Debunking Arguments in Parallel: The Cases of Moral Belief and Theistic Belief” serves as a bridge between the first and the second parts. Baker-Hytch distinguishes four types of EDAs that can be directed against both moral and theistic beliefs and examines how these two kinds of beliefs fare against those different EDAs. By his lights, theistic beliefs are overall less liable than moral beliefs to be undercut by EDAs. For, although both kinds of beliefs fare equally well—or equally poorly—against the companions in guilt argument, theistic beliefs fare better than moral beliefs against the counterfactual argument, the explanatory argument, and the probabilistic argument. Baker-Hytch remarks, however, that this conclusion is drawn on the assumption that the epistemology of moral beliefs is to be considered independently of the epistemology of theistic beliefs. But if one focused on theistic moral belief, and hence considered the possibility of God ensuring the reliability of our moral belief-forming processes, then moral beliefs might inherit the suggested advantages that theistic beliefs have against EDAs.

In “Rationalization, Reasons, and Religion,” Joshua Thurow engages with CSR in defense of the rationality of religious belief. Following a line of argument developed in some of his previous work, he first contends that debunking arguments that appeal to CSR genealogies provide at most partial accounts of religious belief. For they may explain why humans have the inclination to form and sustain that kind of belief but not why individuals or communities hold the particular religious beliefs they hold. In explaining their religious beliefs, believers usually offer reasons—such as cosmological and design arguments, personal religious experiences, and the testimony of trustworthy people—that may well justify those beliefs even if the CSR genealogies are correct. Thurow then considers a countermove by the religious debunker: CSR provides evidence that the reasons offered by believers are nothing but rationalizations inasmuch as their particular religious beliefs are actually the product of epistemically distorting factors. Thurow maintains that this rationalization-based debunking argument only partially undermines the justification of religious beliefs, for theistic worldviews have the resources to deflect to some extent the reasons to think that believers rationalize.

In “The Problem of Natural Nonbelief: Prehistoric Humans, Religious Debunking, and Divine Hiddenness,” Matthew Braddock addresses a
challenge that appeals to a divine hiddenness argument, the cognitive science of religion, the cultural evolution of religion, and the ethnographic record. Succinctly put, the challenge is this: if there is a God who created the human mind, why did millions of prehistoric humans naturally fail to believe in God? In other words, why would God hide from them by endowing them with cognitive capacities that in their environments would only enable them to form non-theistic concepts of highly limited supernatural agents? Proponents of the challenge take prehistoric humans’ natural non-belief in the God of theism as strong evidence that theism is probably false inasmuch as it is much more surprising given theism than given naturalism. Braddock argues that the challenge does not pose a problem for theism both because it is unclear what prehistoric humans believed about gods and because, even if natural non-belief in the God of theism was prevalent among them, it would not actually be very surprising given theism.

In “Milvian Bridges in Science, Religion, and Theology: Debunking Arguments and Cultural Evolution,” Lari Launonen and Aku Visala engage with an EDA against religious belief that appeals to cultural rather than biological evolution. According to this EDA, religious beliefs are unjustified, not because they are generated by biologically shaped cognitive processes that are unreliable as far as those beliefs are concerned, but because they are generated by cultural processes that select for those beliefs for their ability to produce pro-social behavior rather than for their truth sensitivity. Scientific beliefs, by contrast, are truth-sensitive because their cultural fitness depends on their power to produce accurate predictions. Their truth sensitivity explains the great amount of convergence on them that exists across cultures. In response to the EDA in question, Launonen and Visala argue that the difference between science and religion is actually more a matter of degree than a matter of kind, that there is considerable cross-cultural convergence on theistic and Christian beliefs, and that folk Christian beliefs, just as folk scientific ones, are truth-sensitive to the extent that they are constrained by expert beliefs.

The third part of the volume is devoted to EDAs in mathematics. The Benacerraf–Field challenge is a well-known objection to mathematical Platonism. Its proponent argues that, if mathematical entities are, as Platonists claim, mind-independent and causally inert, then we cannot explain how it is that we have cognitive access to the mathematical realm or how it is that our mathematical beliefs are mostly true. As a result, our belief in those mathematical entities is undermined. It has been argued that a similar objection could be raised to robust forms of normative realism, for these maintain that normative facts are independent of our normative beliefs and that they are not natural facts that can enter into causal relations. In “The Epistemological Challenge to Robust Mathematical Platonism: A New Hope?” Mary Leng proposes a novel response to the Benacerraf–Field challenge by drawing on David Enoch’s
“third-factor” defense of robust normative realism against Street’s EDA. According to third-factor responses to EDAs, there is a factor that correlates with both our normative beliefs and the objective normative truths, and that therefore explains their correlation with each other even though those truths are not the cause of our having the normative beliefs we have. In other words, as a matter of fact, our normative beliefs track the objective normative truths even though we did not evolve to track them. Leng maintains that, in the case of arithmetic, logical truths concerning numerosities are the third factor that explains the correlation between our evolutionarily advantageous basic beliefs concerning numbers and the relevant facts about numbers. She also tentatively explains how this account could be extended so as to provide an explanation of our reliability when it comes to mathematical objects other than the natural numbers. The significance of Leng’s response lies in the fact that, unlike other responses to the Benacerraf–Field challenge, it does not intend to meet it by watering down the robustness of the Platonic account of the mathematical realm.

It is a common thought that mathematicians are, for one reason or another, free to make any logical coherent pure mathematical posits they like. If so, then one can reduce mathematical access worries (including worries raised by EDAs against mathematical Platonism) to access worries about a certain kind of knowledge of logical coherence. But how helpful is this move in answering mathematical EDAs? In “What Logical Knowledge is Needed to Account for Our Mathematical Knowledge,” Sharon Berry tries to help readers clarify their views on this question in two ways. First, she reviews and develops a coincidence avoidance framework for evaluating EDAs and proposed solutions to them. Second, she taxonomizes and explains how different intuitive ideas about mathematics lead to different popular positions on what kind of knowledge of logical coherence would be needed to explain our mathematical knowledge in the aforementioned sense.

The fourth and final part of the volume consists of four essays that deal with metaphysical or epistemological matters. In “On Debunking Color Realism,” Daniel Korman and Dustin Locke explore how color-free explanations of color experience are supposed to debunk a robust realist account of color, according to which colors are mind-independent and non-physical properties that are instantiated by physical objects and that supervene on the physical surface properties of their bearers. In their view, such explanations can underwrite a potentially effective debunking argument against robust color realism provided they are supplemented with an evolutionary account of the origins of color vision—such as that according to which trichromatic color vision enhanced our ancestors’ foraging abilities. Even though they examine what it would take for such an EDA to be successful and what resources are available to the robust color realist for resisting it, they refrain from adopting a definite view on
whether the argument is ultimately effective because its cogency hinges on highly controversial issues. This essay will be of interest to epistemologists inasmuch as Korman and Locke offer an in-depth examination of various epistemic principles that may underwrite EDAs in any domain and of how the choice between those principles affects the prospects of debunking robust color realism.

The next three essays are devoted to issues in mainstream epistemology. In “Debunking, Theoretical Indispensability and Irreducible Epistemic Rationality,” Christos Kyriacou maintains that there is a coherent network of basic norms of epistemic rationality that are *prima facie* indispensable for any rational argument. For this reason, if a debunking argument (evolutionary or otherwise) targets those norms, it will fall victim to a *reductio*: it will lead to global radical skepticism or defeat itself or undermine what guides assertion, practical reasoning, and action. Kyriacou calls his maneuver “the Cartesian gambit” because, just as Descartes appealed to a hyperbolic doubt that called into question all of his beliefs in order to demarcate the rational limits of doubt, so too does Kyriacou attempt to call into question the basic norms of epistemic rationality in order to demarcate the rational limits of debunking arguments. Kyriacou further claims that his indispensability argument is stronger than both the Quine–Putnam scientific indispensability argument for mathematical entities and David Enoch’s practical indispensability argument for moral facts.

In “Global Debunking Arguments,” Andrew Moon examines how one should respond to debunking arguments that conclude that one has a defeater for all of one’s beliefs because one believes that they were not formed reliably—as is the case of Alvin Plantinga’s evolutionary argument against naturalism (Plantinga 1993, 2000, 2011). Moon first argues that, even though it is not possible to defeat global defeaters and it is a mistake to think that debunking arguments fail if they are self-defeating, one can nonetheless eliminate a global defeater. This can be done not by reasoning but rather by acting so as to bring about an epistemically good state of affairs. Moon then distinguishes between three types of global debunking argument corresponding to three types of defeater—pure-undercutters, undercutters-because-rebutters, and undercutters-while-rebutters—and considers how one can prevent each type of argument from providing a defeater in the first place. Finally, Moon proposes a solution to the so-called conditionalization problem for Plantinga’s evolutionary argument against naturalism.

In the last essay of the volume, “Global Evolutionary Arguments: Self-Defeat, Circularity, and Skepticism about Reason,” I consider both an EDA that purports to undermine the epistemic justification of the belief in the reliability of our belief-forming processes, and an evolutionary vindicating argument (EVA) that seeks to establish that such a belief is epistemically justified. Whereas the EDA in question falls prey to self-defeat,
the EVA under consideration falls prey to vicious circularity. My interest in those arguments and the problems they face lies in what they might tell us about the possibly aporetic nature of reason. For, if we take the EDA and the EVA in question to consist of true or plausible premises and valid inferences at which we arrive through a meticulous use of reason, then their falling victim to either self-defeat or vicious circularity might be regarded as a sign that, when we push rational reflection on the reliability, or lack thereof, of our belief-forming capacities to the limit, we end up in a situation of aporia from which there seems to be no escape.

Notes

1 See Machuca (2018) for an annotated bibliography on evolutionary debunking arguments in ethics.
2 Bagwell (2021) examines an EDA of this kind and argues that it is self-defeating.

References


