dialectica

dialectica Vol. 69, N° 2 (2015), pp. 185–203
DOI: 10.1111/1746-8361.12098

Evolution and Epistemic Justification

Michael VLERICK and Alex BROADBENT[†]

ABSTRACT

According to the evolutionary sceptic, the fact that our cognitive faculties evolved radically undermines their reliability. A number of evolutionary epistemologists have sought to refute this kind of scepticism. This paper accepts the success of these attempts, yet argues that refuting the evolutionary sceptic is not enough to put any particular domain of beliefs – notably scientific beliefs, which include belief in Darwinian evolution – on a firm footing. The paper thus sets out to contribute to this positive justificatory project, underdeveloped in the literature. In contrast to a 'wholesale' approach, attempting to secure justification for all of our beliefs on the grounds that our belief-forming mechanisms evolved to track truth, we propose a 'piecemeal' approach of assessing the reliability of particular belief-forming mechanisms in particular domains. This stands in contrast to the more familiar attempt to transfer warrant obtained for one domain (e.g., common-sense beliefs) to another (e.g., scientific beliefs) by showing how one is somehow an extension of the other. We offer a naturalist reply to the charge of circularity by appealing to reliabilist work on the problem of induction, notably Peter Lipton's distinction between self-certifying and non-self-certifying inductive arguments. We show how, for scientific beliefs, a non-self-certifying argument might be made for the reliability of our cognitive faculties in that domain. We call this strategy *Humean Bootstrapping*.

1. *Introduction*

We have evolved, and so have our minds. But what have our minds evolved to do? It would be nice if we could answer that they have evolved to produce true beliefs, at least in certain domains where we like to think that our beliefs are true. Notable among these is the domain that the theory of evolution itself occupies, the domain of scientific beliefs. But according to the evolutionary sceptic, the fact that our belief-forming mechanisms evolved means that we are not warranted in thinking that *any* of our beliefs are true. They evolved for fitness, and fitness is not truth.

This sceptical doubt is as old as the Darwinian theory of evolution itself, since Charles Darwin formulated it:

> 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 1881)

A growing number of naturalists, however, have put up a strong case against evolutionary sceptical worries (Fales 1996; McKay and Dennett 2009; Law 2012;

[†] Department of Philosophy, University of Johannesburg, PO Box 524, Auckland Park 2006 (South-Africa). Email: michaelvlerick@gmail.com; abbroadbent@uj.ac.za

Wilkins and Griffiths 2012; Boudry and Vlerick 2014). They have forcefully argued that evolutionary considerations are not at odds with epistemic reliability. In fact, as McKay and Dennett put it, "although survival is the only hard currency of natural selection, the exchange rate with truth is likely to be fair in most circumstances" (McKay and Dennett 2009, 509). In other words, in most cases our cognitive faculties can be expected to have evolved for truth-tracking. As one of the authors has put it in another paper, "evolution does care about truth" (Boudry and Vlerick 2014).

We think these defences against evolutionary scepticism are convincing. We are less confident, however, that they warrant a strong confidence in all or even most domains of beliefs. It is easy to suppose, as Descartes did, that refuting scepticism amounts to justification: to putting knowledge on a secure foundation. But, at least in the case of evolutionary scepticism, this does not follow. The evolutionary sceptic doubts *all* of our beliefs; the defence merely establishes that this sceptical worry, concerning *all* of our beliefs, is not warranted. Fitness and truth are not in universal tension; indeed, there are good reasons to think that the ability to form beliefs about a given domain is selected for only if the beliefs are largely true. But this is not enough to show that all of our beliefs are true (thankfully! since that would be incredible). Less obviously, nor is it enough to show that in *any particular* domain of beliefs, our beliefs are likely to be true, beyond domains on which selection pressure operated directly. In particular, it is not enough to show that the domain of beliefs where the theory of evolution resides, the domain of scientific beliefs, is likely to contain true beliefs, given that it has been produced by evolved cognitive faculties.

There is thus an important distinction to be made between responding to the evolutionary sceptic (which we call the *negative project*), and providing our beliefs in any given domain – for example, our scientific beliefs – with epistemic justification (which we call the *positive project*). Our central aim in this paper is to contribute to the positive project.

The negative project has received a lot of attention in the literature. It has been convincingly argued, contra Plantinga (1993) and other evolutionary sceptics, that scepticism is not warranted merely by the fact that our cognitive faculties evolved. Grounding the origin of our cognitive faculties in an evolutionary process does not undermine the reliability of those faculties. We take this to have been convincingly established in the literature. We summarize, but do not seek to contribute to the negative project.

The positive project of providing our best theories about the world with justification is underdeveloped in the literature (cf. De Cruz et al. 2011). Of course, naturalists have made positive claims: Wilkins and Griffiths (2012), for example, suggest that common sense and scientific beliefs can be justified. However, the argument is not made out in any detail: it is vulnerable to uncharitable readings on which it is simply fallacious, and more charitable readings require additional material that is not supplied in any work of which we are aware. It is here that we hope to contribute something useful.

In section 2, we summarize the negative project, relying largely on a good recent implementation in two papers by Wilkins and Griffiths (Wilkins and Griffiths, 2012; Griffiths and Wilkins in press). In section 3, we distinguish the positive from the negative project. We distinguish *wholesale* from *piecemeal* scepticism and justification, and show that refuting wholesale scepticism is insufficient for establishing piecemeal justification. In other words, refuting evolutionary scepticism in general, as Wilkins and Griffiths do, is not enough to establish that any particular domain of beliefs is justified.

In section 4, we embark on the positive project, setting out the elements of an alternative strategy that we call *Humean Bootstrapping*. This strategy draws on externalist responses to other kinds of scepticism, notably the response of reliabilists such as Peter Lipton in the context of induction. The core idea is that there is more than one kind of circularity, and that some circles are better than others. In certain domains, we argue, the use of cognitive faculties to justify the claim that those cognitive faculties are truth-producing amounts to a virtuous circle, not a vicious one. We apply Lipton's test: if the hypothesis were false, the evidence would have been different; or, as he puts it even more succinctly, "if the conclusion had not been true, you would have noticed" (Lipton 2000, 184–185). It is a contingent fact that a given faculty's outputs satisfy this test, and the assertion of that fact is subject to empirical justification; the strategy is thus a piecemeal one, which must proceed faculty by faculty. We cannot connect the fact that our cognitive faculties have evolved to their wholesale reliability; but we can justify the claim that particular faculties and methods are reliable in particular domains. We can establish that they track truth *piecemeal*, by establishing that if they did not, we would have noticed. We argue that this is true of common-sense beliefs and of those in the scientific domain. Section 5 deals with the cognitive underpinning of our proposed strategy for epistemic justification, and section 6 concludes.

2. *The negative project*

Building upon previous research in evolutionary epistemology, Wilkins and Griffiths (2012; Griffiths and Wilkins in press) undercut evolutionary sceptical worries by showing that truth-tracking and fitness-tracking are not alternatives. They are, according to Wilkins and Griffiths, different levels of explanation, not potentially rival explanations. It makes no sense to ask whether the bird's wing evolved for locomotion or 'merely' to enhance the bird's fitness, because locomotion is part of what the wing contributes to a bird's fitness. Likewise, it makes no sense to ask whether squirrels evolved long bushy tails to enhance balance, or 'merely' to enhance fitness. And it makes no sense to ask whether the human mind evolved for truth-tracking or 'merely' for fitness enhancing. Truth-tracking is how belief-forming mechanisms enhance fitness, just as balance is for squirrels' tails, and locomotion for birds' wings. Truth, they conclude, is 'the currency of evolutionary success in the domain of cognition' (Griffiths and Wilkins in press).

Nevertheless, empirical evidence abounds for the existence of stubborn cognitive biases. Humans err. They do so systematically, and in diverse areas. Not only do our 'fast and frugal' heuristics often lead us astray in areas ranging from probability calculus to logic (for an extensive overview see Kahneman 2011), but our reasoning also seems to be systematically biased when making judgements under uncertainty, when the costs associated with false positive errors (asserting something is the case when in fact it is not) and false negative errors (asserting something is not the case when in fact it is) have been asymmetric over evolutionary history. This prompts Haselton and Nettle (2006) to dub the bias-riddled human subject "the paranoid optimist".

This, Wilkins and Griffiths reply, does not mean truth and fitness are at odds with each other. It merely shows that there are constraints to truth-tracking. First of all, organisms are subject to resource constraints. Cognition is very costly. Human brains account for up to 20 per cent of oxygen consumption.[1] Perfect truth-tracking is simply not affordable. The expensive nature of the cognitive apparatus constrains the ability to track truth. There is an undeniable trade-off between accuracy and cost, and when the added value of increasing accuracy does not weigh up against the added costs of such cognitive improvement, it will not be selected for.

Not only that, but the attempt to maximize accuracy given cost constraints may actually give rise to fallacies and biases. The tendency to ignore base rates, for example, might lead to error in certain situations, but it may lead to more truth in situations where ignoring the base rate makes little difference. It will do this because each operation involves a cost, and ignoring base rates saves costs, meaning that a larger number of such operations may be performed given a certain resource base. Considering that base rate calculations are largely superfluous in natural environments since these environments rarely present us with random

[1] This, Wilkins and Griffiths note, refutes another debunking line of argument, claiming that truth-tracking might be adaptive, but that human cognition just is not an adapted feature, but a mere frivolous 'spandrel' or exaptation. As Wilkins and Griffiths point out, following Fales (1996, 440–441) among others, such an expensive set of faculties would have been selected away if it had not conferred our ancestors with an adaptive advantage.

sampling (Gigerenzer 1991), we can see why such a bias evolved. Thus a tendency to commit a fallacy may evolve, not merely as a trade-off between cost and accuracy, but in order to *improve* accuracy given cost constraints. Cheap fallacy can, in the right circumstances, deliver more truth than costly validity.

Moreover, Wilkins and Griffiths continue, another constraint arises from the "intrinsic logical structure of many cognitive tasks". When an organism makes a decision under uncertainty, given fixed resources, it is logically impossible to reduce the risk of one type of error (let's say false negatives) without increasing the risk of the opposite type of error (in this case false positives). Natural selection, therefore, faces the task of achieving an optimal balance between the two types of errors. It is wrong to argue from the existence of cognitive fallacies that evolution does not care about truth or worse, selects against truth, since these fallacies often aim at *maximizing* truth-tracking abilities within the confines of both cost constraints and constraints arising from the logical structure of recurring problems in the environment. The significance of this line of argument is not to show that any particular domain of belief is free of error. Rather, it shows that the exhibition of a cognitive fallacy is not enough to establish that evolution does not select for truth. Cognitive biases, the argument states, are actually truth-maximizing given constraints.

This, however, does not establish that our evolved faculties *are* truth-producing in any particular domain. Nor does it resolve the worry that the modern environment differs importantly from the environment that shaped these faculties. Establishing these things, we take it, is part of the positive project of showing that particular domains of belief – sets of beliefs that we actually hold – have been formed by mechanisms that are probably reliable.

3. *From negative to positive?*

It is easy to suppose that if the sceptic is refuted, or at least fended off, then the task of justifying those beliefs we wish to believe is done. The sceptic doubts; the reply shows that the doubt is not warranted; what further cause is there to worry? But this is not right – at least, not in the context of evolutionary scepticism.

Evolutionary scepticism is (typically) *wholesale*: it undermines all of our beliefs, by arguing that the biological structures that produced those beliefs were subject to selection pressures that had nothing to do with truth. The reply is that truth is not an alternative to fitness. Truth-tracking is the contribution that cognitive fallacies make to fitness, just as flight and balance are the contributions to fitness made by birds' wings and squirrels' tails respectively. The reply, therefore, is a negation of wholesale scepticism. Negating the claim, however, that all our belief-forming mechanisms are unreliable does not entail that *any* *particular subset* is reliable. Perhaps this simple logical point is obvious, but it is not always prominent in the literature.

There are some positive piecemeal justifications on offer for common-sense belief. For example, if we accept that truth is the way that cognition confers a selective benefit, then it is plausible that, in the domain of common-sense beliefs, organisms that have beliefs track truth so as to obtain "as much relevant truth as they can afford" (Wilkins and Griffiths 2012, 137–138). Common-sense beliefs, therefore, meaning "everyday beliefs guiding our mundane action", are produced by cognitive adaptations that track truth (Ib.). Given that the differences between contemporary and ancestral environments are not so great that common-sense beliefs are likely to misfire radically in modern environments, we can use this evolutionary argument to provide core common-sense beliefs (such as the belief in an external world and the belief in other minds) with justification (cf. Stewart-Williams 2005).

But as Wilkins and Griffiths point out, the interesting project in the context of epistemic justification does not so much concern our day-to-day, common-sense inferences about our immediate environment, but our elaborate scientific explanations aimed at tracking truths that transcend our particular 'Umwelt'[2] (Wilkins and Griffiths 2012, 140). The main task awaiting the naturalist is not so much debunking the evolutionary sceptic – who denies us justification of even our most basic and self-evident 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 (the positive project). Too often a sharp distinction is not drawn between these two projects, and it is tacitly assumed that rebutting the sceptic by establishing that natural selection shaped our cognitive faculties for truth-tracking, yields us the positive project at the same time. This tempting, but fallacious, line of thought states that our scientific beliefs are the product of evolved faculties, and because these faculties evolved to track truth in ecologically relevant contexts, they can be expected to track truth *in general*. Therefore, the argument goes, the beliefs that they give rise to in the scientific domain are justified. This strategy we call the "evolutionary wholesale strategy" for epistemic justification. Although it is tempting, on closer inspection it harbours two major flaws.

The argument may be refined as follows:

(1) Our cognitive faculties are truth-tracking (premise)
(2) Scientific beliefs are produced by our cognitive faculties (premise)
(3) Scientific beliefs track truth (conclusion)

[2] This term coined by Von Uexküll (1909), designates the particular realm of awareness in which every species is encapsulated, as the outcome of its perceptual and conceptual categories.

The argument, as stated, is valid. The first major problem, however, concerns premise 1. As Boulter (2007) points out, only common-sense beliefs – the type of beliefs that enabled our ancestors to cope with their environment – are on the radar of natural selection. These beliefs include belief in the existence of other minds, the occurrence of past events, and the reliability of perception (De Cruz et al. 2011, 520). In other words, our cognitive faculties evolved to track truth in a narrow domain of ecologically relevant beliefs. They did not evolve to track truth *in general*. At most, given that relevant environmental circumstances have not changed too radically since the ancestral environment in which our faculties evolved, evolutionary considerations give us some justification for our core common-sense everyday beliefs – not a wholesale justification of all of the output of our cognitive faculties.

The second problem with the argument is of a general epistemological kind. Even if Premise 1 could be justified as it stands, the argument would be ineffective. For what could possibly warrant Premise 1, other than scientific evidence for the reliability of our cognitive faculties? But if we accept such evidence, we implicitly accept 3, the conclusion. If we reject 3, we will reject the evidence adduced for 1. As stated, the argument is not circular; but if the justification for Premise 1 were included, then 3, or something very much like it, would appear. Fully stated, the argument would be premise-circular; the concise statement above conceals this only by suppressing all the justificatory steps underlying Premise 1.

For this reason, there is no reasonable person who will accept the first premise who has not already accepted, albeit implicitly, the conclusion that our scientific beliefs are generally true. Anyone doubting the conclusion, on the other hand, will also doubt the premise. Thus the argument is epistemologically mute. The conclusion cannot be epistemologically significant given that the epistemological stance it states is one that an audience must already adopt in accepting the premises.

While any naturalized epistemology – using instances of knowledge to make claims about knowledge in general – must learn to live with circular arguments, circular arguments are not all alike. Some are viciously circular, or – like the argument above – reveal themselves as being viciously circular when the justification for their premises is explored. They do absolutely nothing for us, epistemically speaking. They provide neither a reason to believe something we did not previously believe, nor any further warrant for beliefs we already have. The recipient of a viciously circular argument is epistemically in just the same position as she was beforehand.

Not all circular arguments are viciously circular, though. 'Virtuously circular' arguments are like viciously circular arguments in that they are unable to convince the sceptic who doubts any given premise. Nevertheless, for those who accept the premises, they provide further reasons for believing what one already believed. In the next section we seek to characterize virtuous circularity more explicitly, following Peter Lipton's treatment. Our point here is that the wholesale evolutionary argument for epistemic justification is not virtuous in any sense. There is nobody who will accept the premises who has not already accepted the conclusion; and conversely, anyone doubting the conclusion will also doubt the premises.

One might draw back from a wholesale argument of the very bold kind previously advanced, but still seek to build on the success of cognitive mechanisms in forming true common-sense beliefs in order to justify scientific beliefs. For example:

1′. Our cognitive faculties are truth-tracking concerning mundane, everyday matters. (premise)
2′. Faculties that are truth-tracking concerning mundane, everyday matters will be truth tracking concerning scientific matters. (premise)
3′. Scientific beliefs are produced by these cognitive faculties that are truth-tracking concerning scientific matters. (premise)
4′. Scientific beliefs track the truth. (from 1–3)

This argument is piecemeal in the sense that it focuses on scientific beliefs, but it shares the problematic characteristic of wholesale approaches, namely, what we might call *domain-switching*. The argument attempts to employ the fact we have good reason to accept that our faculties produce true beliefs in one domain (here, common sense) in order to show that they produce true beliefs in another domain (here, science). Perhaps that would be a good strategy if strong logical links existed between the two domains. But for two domains that are not strongly connected, it is very hard to see how it can ever succeed.

Thus the difficulty with the argument just set out is Premise 2′. The complex representations of modern scientific theories, we argue, are by no means an extension of common-sense beliefs. Given that natural selection did not shape our minds to track scientific truths about the world, and that modern science – as De Cruz and De Smedt (2012) point out – is a by-product of cognitive faculties that evolved for survival and reproduction, justification cannot merely be extended from the domain of common-sense beliefs to the domain of scientific beliefs. In fact, evolutionary considerations have been used to simultaneously justify core common-sense beliefs and debunk scientific beliefs (cf. Stewart-Williams 2005).[3]

[3] Stewart-Williams (2005) argues that the fact of human evolution simultaneously undermines both radical scepticism and reasonable confidence in the accuracy of our less basic and more complex representations. Exploring the question whether the fact that certain components of our worldview have an evolutionary origin implies that these aspects accurately depict the world, he leans towards a negative answer (Ib., 792). Nevertheless, he finds ground in evolutionary considerations to justify core common-sense beliefs against radical scepticism and solipsism. More precisely, he argues that given that our belief in a mind-independent, external world is rooted in innate aspects of the mind (since it is not strictly deducible from sensory information) and that evolution by natural selection shaped these aspects of the mind, such a world must indeed exist (Ib., 794).

In order to see why justification cannot be unproblematically extended from the domain of common-sense to scientific beliefs, compare our 'folk sciences' to their proper scientific counterparts. 'Folk sciences' refer to our common sense (or, perhaps, Stone Age) beliefs and inference patterns shaped to deal with ecologically relevant aspects of the world, such as for instance the world of objects (folk physics) and living organisms (folk biology). Both of these folk theories are innately determined (i.e., we are predisposed to develop these beliefs – see, for instance, Spelke (1991) and Baillargeon (1991) on innate folk physics and Atran (1998) on innate folk biology – and enable us to deal with the environment in a survival and reproduction promoting way. They do so, as explained when setting out the negative project (see section 2), by tracking affordable and relevant truth in those domains.

Folk physics, for instance, enables us to gauge fast and accurately where a thrown stone will land, how long a falling rock takes to hit the ground or what will happen when a moving object collides with an object in rest. Folk biology, on the other hand, enables us to predict the behaviour of unknown animals successfully based on their resemblance with known species, gauge whether an unknown plant is edible based on resemblance with known plants and in which environment certain organisms will thrive. Improvements in accuracy will be selected for to the extent that they are worth their cost, and to this extent, evolutionary pressure improves folk science's ability to track true states of affair.

However, this does not imply that those folk sciences provide us with accurate or even approximately accurate representations of the world *in general* – i.e. outside our ecological niche. In fact, science proper suggests that they are notoriously inaccurate. Folk physics, for instance, is radically contradicted by Newtonian and a fortiori Einsteinian physics. It ascribes something like an 'impetus' to moving objects and assumes that every object's natural state is at rest. This however, as Pinker (1997, 321) points out, is not surprising. In the real world, Newton's laws are masked by friction (from the air and contact with the ground). Our folk physical beliefs serve us well in friction-ridden dealings with medium-sized dry goods. Indeed, they may represent the best way to maximize truth in a resource-restricted context. Similarly, folk biology with its classification of the organic world into a complex taxonomy, based on an intuition of hidden traits or essences that members of each group at each level share with each other (Atran 1998; Pinker 1997, 323) is useful but flatly contradicted by evolutionary theory and population thinking.

The cognitive faculties and processes underlying our common-sense beliefs, it should be clear, did not evolve to track biological and physical truth in general, only physical and biological truth in a very specific ecological situation. Any evolutionary argument extending epistemic justification to scientific beliefs from the premise that our cognitive faculties evolved to track truth, overlooks both the depth and the scope at which evolved truth-tracking mechanisms actually track truths. Selection pressures ensure nothing more than that common sense tracks ‘shallow’ truths – the kind of truths that promote survival (e.g., whether plants are poisonous or edible, animals threatening or docile, etc.) – in the ‘narrow’ ecological niche in which we evolved. Not the kind of truths modern science aims at in its attempt to uncover fundamental laws underlying empirical data which stretch out far beyond our ancestral ecological niche. Extending justification to the domain of scientific beliefs, therefore, misfires.

In fact, if anything, the cognitive faculties and processes underlying our evolved common-sense modes of understanding form a barrier to overcome when attempting to track truth in the enriched conceptual framework in which science operates. In an interesting paper, De Cruz and De Smedt (2007) show how our evolved modes of understanding the world systematically distort scientific discourse. Science, it appears, is more often than not at odds with common sense – or at least good science is. According to Wolpert: “if something fits with common sense it almost certainly isn’t science” (Wolpert 1992, 11). Any attempt to justify our elaborate scientific representations by arguing that they are underwritten by cognitive faculties and processes leading to truth-tracking common-sense beliefs, gets it wrong. There is no hope of doing with evolutionary arguments what Descartes (1641) failed to do with God. There is however another strategy available to the naturalist. In contrast to the wholesale approach to justification outlined above, this strategy takes a piecemeal approach. We call it the ‘Humean Bootstrapping’ strategy.

4. *Humean Bootstrapping*

The received epistemological stance of the naturalist is that our belief-forming faculties are reasonably reliable. This stance is non-optional: were our faculties completely unreliable, any attempt at justifying our beliefs is doomed to fail. It is, therefore, the default assumption and the working assumption from which all arguments are mounted. We have rejected the strategy of transferring warrant from one domain to another – of showing, for example, that our cognitive faculties are truth-sensitive regarding mundane, everyday matters, and then seeking to argue on this basis that they are also truth-sensitive regarding scientific matters. Instead, we seek direct justification for the reliability of our cognitive faculties in scientific domains, by appealing to the default, working assumption of reliability, and ‘bootstrapping’ our way from there.

How could any such strategy hope to avoid vicious circularity? In an important and rich paper, Peter Lipton sets out to explain how the inductive justification of induction may be worthwhile even if it is not a reply to a sceptic about induction

(Lipton 2000). Lipton is not the first to suggest this, but his treatment is among the clearest and most thorough, and his concern with assessing the warrant for scientific beliefs makes it especially relevant here. Along the way he offers characterizations of circularity in general, solutions to raven and grue paradoxes, and rebuttals of the no-miracles argument for realism and the pessimistic meta-induction against it. We are here concerned only with his general strategy for rehabilitating the inductive justification of induction.

Lipton points out that there are prima facie reasons to take the inductive justification of induction seriously, notwithstanding its bad reputation. In particular, we single out the following reason:

> the claim that an inductive method is reliable is tantamount to a physical hypothesis that there is a correlation between the output of the method and the state of the world. If there can be legitimate inductive arguments concerning the expansion of metals when heated, there can be legitimate inductive arguments about the reliability of thermometers. (Ib., 180).

Lipton's strategy is to show that the inductive justification is no worse than inductive arguments generally: inductive arguments for physical hypotheses, for example. If he succeeds, this means that the inductive justification of induction will have value for someone who is prepared to rely on inductive inference. Of course, we all fall into that category whether we like it or not. The question, then, is whether he can indeed show that the inductive justification of induction stands or falls alongside inductive inferences generally: whether he can show that the circularity of the inductive justification of induction does not matter.

Lipton's argument turns on the distinction between two inductive assessments of inductive methods: those that are self-certifying and those that are not. A self-certifying assessment of an inductive method is such that the use of the method ensures that the assessment is positive. For example, consider the claim that a particular thermometer is generally reliable. The inductive method under assessment is the use of this thermometer to measure temperatures. Compare two ways of assessing this method. One way would be to appeal to previous readings of that very thermometer. This would be self-certifying (or practically so), because there is (practically) no way for the track record of the thermometer to reveal occasions on which it has incorrectly measured the temperature. Another way would be to appeal to other readings of independent thermometers, and correlate them with readings of the thermometer under assessment. This would not be self-certifying, because discrepancies between readings would provide information that could reveal, or at least point us to, cases where the thermometer has given erroneous readings. Lipton captures the difference in terms of a Nozickian tracking requirement (Nozick 1981). In the second case, "if the conclusion had not been true, we would not have made the inference" (Lipton 2000, 184); whereas in the first case, we might have.

Lipton's claim is that the inductive justification of induction is not necessarily self-certifying. He is careful not to make his claim for just *any* inductive assessment of induction, since he leaves space for there to be good and bad such assessments, just as there can be good and bad inductive arguments more generally. (He is sensitive to the problem that we above characterized as domain-switching.) However, his general point is that we ought not to write off a given inductive justification of a given inductive method merely because the justification, like the method it justifies, is inductive. Unlike premise-circularity, rule-circularity (where the rule used to arrive at a conclusion is stated by that conclusion) does not guarantee that the conclusion is warranted by the premises. The success of an attempt to inductively justify an inductive method is not a foregone conclusion.

The logical point is, we think, sound. If a method is treated as reliable in the first place, there is no special reason to restrict its use so as to exclude the self-assessment of that method. Such self-assessment need not be self-certifying. Whether the resulting cognitive endeavour is a good one or not is a question to be answered in the same way that one answers these questions generally. But it can be good. When it is, it is reasonable for us to regard ourselves as having further reason to accept the reliability of M, despite the fact that our reason depends on the use of M. This situation we call Humean Bootstrapping: Humean because it begins with the fact that we cannot but use certain methods (Hume 1748); and bootstrapping because it involves the use of methods to provide reason for accepting the reliability of those very methods.

Humean Bootstrapping, in this regard, provides us with a framework in which the initial assumption of the reliability of our faculties (an assumption – as pointed out – which is non-optional) can be strengthened, not by securing a link between their origin in an evolutionary process and their reliability – as in the case of the wholesale evolutionary argument – but by gathering new evidence that these faculties are indeed reliable.[4] Whereas it is circular and self-certifying to argue that our beliefs are justified because our cognitive faculties evolved for truth tracking, it is circular but not self-certifying to use our own cognitive faculties to assess their reliability. Indeed, more than often such assessments have brought to light that (some) of our faculties and inference-mechanisms are actually not as reliable as we assumed (Kahneman 2011).

Consequently, rejecting wholesale positive justificatory arguments need not mean giving up on the project of gaining warrant for our beliefs. We must simply

[4] In this regard, our approach to epistemic justification is similar to Goldman's (1975) process reliabilism in which a belief is justified if the (causal) process that produced the belief is a reliable one.

do so piecemeal, using our abilities in given domains to assess those abilities. Provided that this is not self-certifying, our procedure is not epistemically impotent. It may not move someone who is so sceptical that she refuses to grant any kind of reliability to our cognitive faculties at the outset. But it can help someone who rejects such a sceptical position and is willing to remain in "the circle of knowledge" (Russell 1912) to see how it is possible to gain confidence in the output of these faculties. In short, epistemic justification of our representations cannot be grounded wholesale – in Cartesian fashion – in their evolutionary origin, but there is a Humean alternative available to the naturalistic epistemologist. In implementing this strategy, we start with the assumption that our faculties are overall reliable (at least reliable enough to assess their own reliability) and seek to increase our confidence by putting those faculties to the test, or – in Lipton's lingo – by tracking their track records.

5. *Humean Bootstrapping for humans?*

The question remains, can we in fact use our faculties to reflect on their reliability? Inductive methods in general are not necessarily self-certifying; but it remains to be established that in any given domain, our faculties are in fact capable of performing any meaningful assessment of their own reliability. How can we adopt a critical stance towards our own cognitive output, in the context of scientific inquiry? From a naturalist perspective, we need more than just a demonstration of the logical possibility of such a manoeuvre: we need a naturalistic explanation for this remarkable ability which underlies our ability to seek justification for our beliefs. In order to do so, we turn to the cognitive sciences.

One way to construct an answer is as follows. Evolution, according to many cognitive scientists and evolutionary psychologists (e.g., Fodor 1983; Tooby and Cosmides 1992; Mithen 1996; Boyer 2000; Carruthers 2006) has endowed us with a number of content-rich, domain-specific cognitive systems, often called modules.[5] These innate cognitive systems evolved to enable us to deal with important recurrent aspects of our environment, such as other people, inanimate objects and living organisms. They underlie our "intuitive ontologies" as Boyer refers to them, providing us with "a series of category-specific intuitive principles that constitute an evolved 'natural metaphysics'" (Boyer 2000, 277). In other words, they underlie our intuitive thinking about the physical, social and natural world – i.e., our folk physics, folk psychology, folk biology. As pointed out

[5] We are not wedded to the 'modularity' thesis, which is an empirical thesis. Whether this thesis is consistent with empirical facts about our cognitive make-up, is a matter of psychology. We just select the modularity thesis to demonstrate one way of spelling out, naturalistically, how Humean Bootstrapping might occur in humans.

previously (section 3), these folk representations endow us with characteristically shallow and narrow truth-tracking. Therefore, we have argued, these cannot form the basis of an argument for the general reliability of scientific beliefs.

Evolution, however, also equipped us with the cognitive ability to override these 'modular' outputs and engage in critical reflection (Vlerick 2012). How else could we reflect on these folk theories as tracking truth on a shallow and narrow level? This ability to override the belief output of our domain-specific, content-rich cognitive systems, is itself an evolved cognitive ability, and one that is the hallmark of human intelligence. As Fodor (1985, 4) puts it, "what is most characteristic, and puzzling about the higher cognitive mind, [is] its non-encapsulation, its creativity, its holism and its passion for the analogical." In other words, rather than its content-rich domain-specific cognition, the human mind's distinctive feature is precisely its ability to reason outside the confines of its various 'intuitive ontologies'.

A number of cognitive scientists and psychologists have addressed this remarkable aspect of human cognition. According to Carey and Spelke (1994), overriding the output of our innate cognitive systems is the result of "mapping across domains" (Ib., 180). This happens when the core principles of one knowledge system are applied to the set of entities of another system, thereby escaping the principles that naturally – i.e. in virtue of our nature – fit these entities. By devising and using systems of measurement in physics, for instance, scientists create a 'mapping' between the core knowledge system of numeracy and that of physics. Therefore, the principles governing the behaviour of physical bodies are no longer those of cohesion, continuity and contact – the innate principles underlying our folk physical representations (see Spelke 1991) – but the core principles of the system of numeracy, such as 1-to-1 correspondence, succession and the like.

Similarly, according to Carruthers (2006) the human mind – which he characterizes as "massively modular" – is endowed with two reasoning systems. The first corresponds to the processing of the modules: it is arranged in parallel and operates swiftly and unconsciously. The second supervenes on the activity of those systems: it is realised by mental rehearsal in general and inner speech in particular and operates more slowly and consciously. It integrates, in other words, the content outputs of the various modules, overriding the results of the first system (Ib., 254).

Natural language, which according to Pinker (2007) provides us with a window into human thinking and thought processes, is gorged with tangible evidence of this remarkable cognitive process. As Lakoff and Johnson (1980) have pointed out, metaphor or analogy are pervasive in language. Spatial metaphors, for instance, such as up and down, front and back, in and out, etc. organize entire systems of abstract concepts (e.g., numbers go up and down, and so do feelings, moral standards, professional careers, etc.), These analogies show that the human mind is very prone and apt to apply elements from one domain of cognition to a different domain (such as applying spatial reasoning to numerical values). Metaphors in language, in other words, are proof of the way our mind co-opts reasoning patterns that are grounded in a particular innate knowledge system to extend it to different domains.

This distinctive cognitive ability – i.e., the ability of the human mind to override the content of the intuitive ontologies it holds in virtue of its content-rich modules evolved to deal with recurring and ecologically relevant problems – literally opens up our perspective on the world. We are no longer encapsulated in a particular Umwelt or an "evolved natural metaphysics" (Boyer 2000) as other species are, but represent the world in ways that transcend this given perspective both in depth and in scope. The driving force behind this ability is our possession of what one of the authors has called an "epistemic orientation" (Vlerick 2012), i.e., an epistemic goal (truth) and epistemic values or criteria for realizing this goal – such as coherence, predictive accuracy, scope and simplicity.

This provides us with the necessary compass to track truth beyond our cognitive niche and functions as an engine for epistemic improvement. We are able to take the outputs of our belief-forming mechanisms, assess them in terms of those epistemic criteria, and possibly reject them. We may not be able to do this instantaneously or easily; and we may not be able to rid ourselves of cognitive tendencies that we recognize as sub-optimal. Nonetheless, we are able to improve our considered beliefs, and direct our considered actions, according to the outcome of a process of reflection, as exemplified by science.

Rather than being endemic to modern science, this epistemic orientation appears to be an integral part of our cognitive endowment. According to Papineau (2000, 202), the search for truth is an innate drive, much like hunger and the desire for sex. Carruthers (2006, 347) argues that the epistemic values or criteria guiding our search for truth are common to all humankind, from modern scientific circles to illiterate hunter-gatherer societies. Evolution, in this regard, provided us both with a set of fitness-boosting, content-rich representational frameworks (giving rise to our 'default' common sense outlook on the world) and the cognitive tools and compass to actively reflect on these beliefs – and ultimately on the reliability of the faculties producing these beliefs.

A likely evolutionary explanation for this ability is that it evolved in adaptation to environmental variability. As Sterelny (2003, 170) argues with Potts (1996), the physical environment of hominid evolution has been remarkably unstable. In response, it seems plausible, our cognitive apparatus evolved the ability track truth in a wide variety of different contexts (consider the need to detect danger such as predators and poisonous plants and track prey animals in environments ranging from the Antarctic to the Amazonian rain forest). This could account for the fact that we evolved an epistemic compass and a characteristic drive to investigate and actively track truth in out surroundings, rather than merely rely on content-rich modules attuned by natural selection to a fixed set of relevant aspects of our environment.

Whatever the explanation for its origin, however, our ability to transcend our Umwelt combined with this evolved 'epistemic compass' explains why science is able to part ways with intuition, and why it is able to track truth in such a foreign and "enriched conceptual framework" (Wilkins and Griffiths 2012, 140). It also explains why, in Wilkins and Griffiths' words, we are able "to use those [cognitive] faculties to debug themselves". It is, therefore, in this ability or cognitive process, we argue, that the roots of epistemic justification in general, and justification of our scientific representations in particular, need to be located. Not in our 'given' common sense worldview (attuned by natural selection to a set of relevant environmental properties), but – quite to the contrary – in the cognitive processes enabling us to reflect critically on this worldview, expanding it and overthrowing it where necessary.

Note that grounding the epistemic process by which we can bootstrap our way to justified (scientific) beliefs in a natural, evolutionary process does not commit the fallacy exposed in this paper; namely, that of deriving epistemic justification in the domain of scientific beliefs by grounding them in evolved – and therefore supposedly truth-tracking – cognitive faculties. The motor of epistemic justification resides in the process of 'tracking our track records'. The (tentative) cognitive and evolutionary story of how this process is realized in human brains and what selective conditions might explain its evolution, is only developed to show how our approach can be accounted for within a naturalist framework (i.e., without invoking or presupposing abilities that cannot be accounted for scientifically). In short, scientific beliefs are not justified because they are rooted in our (evolved) cognitive ability to override modular output. They are justified to the extent that they stand up to critical and reflective testing of the reliability of the cognitive processes underlying their formation. The human ability to engage in this bootstrapping process, however, is – if the general tenor of this section is on the right track – rooted in the remarkable human cognitive ability to override modular output.

6. *Conclusion*

In response to evolutionary sceptics who argue that given their origin in a natural evolutionary process we should not accord any reliability to the output of our cognitive faculties, a growing number of evolutionary epistemologists have made a strong, empirically supported case that natural selection shaped our cognition to track truth. This negative project of rebutting the sceptic, however, should not be conflated with the positive project of justifying our beliefs in any given domain, notably our scientific beliefs. Failing to draw a sharp distinction between these two projects can lead to the problematic claim that evolutionary considerations warrant justification of complex scientific theories. This, we have argued, is false.

Instead, we offer an alternative strategy available to the naturalist: Humean Bootstrapping. Rather than attempting to find foundations for our belief-forming mechanisms, wholesale, in their evolutionary origins, the naturalist should assess the reliability of these faculties empirically and in a piecemeal fashion. That is, per faculty and per domain in which the faculty is employed. Drawing on recent work in the cognitive sciences, we have attempted to show how the cognitive ability to engage in such a justificatory process might be accounted for by the naturalist. We have argued that the ability to reflect critically on the output of our cognitive faculties is precisely what enables us to track truth in the enriched conceptual framework in which science operates, and that it is in this remarkable ability that we should anchor epistemic justification.

We think that whether our cognitive faculties are reliable in any given domain, notably the domain of scientific inquiries, depends crucially on the empirical question of how they perform, and cannot be settled by 'borrowing' warrant from the reliability of those same faculties in the domain of everyday, mundane beliefs. We hope that we have outlined a positive epistemological project to be pursued by the philosopher of science. This approach of tracking our track-records could also, we think, prove to be very relevant in the scientific realism debate.

Finally, although our focus has been on the positive project, we hope also to have offered the naturalist a better defence against the sceptic and the traditional foundationalist, who often are very keen to point out the circularity of any naturalistic approach.

References

Atran, S. 1998, "Folk Biology and the Anthropology of Science: Cognitive Universals and Cultural Particulars", *Behavioral and Brain Sciences*, **21**, pp. 547–609.

Baillargeon, R. 1991, "Physical Reasoning in Infancy", in: M. Gazzaniga, ed., *The Cognitive Neurosciences*, Cambridge, MA: MIT Press, pp. 181–204.

Boudry, M. and Vlerick, M. 2014, "Natural Selection Does Care About Truth", *International Studies in the Philosophy of Science*, **28**, 1, pp. 65–77.

Boulter, S. 2007, "The "Evolutionary Argument" and the Metaphilosophy of Commonsense", *Biology and Philosophy*, **22**, pp. 369–382.

Boyer, P. 2000, "Natural Epistemology or Evolved Metaphysics? Developmental Evidence for Early-Developed, Intuitive, Category-Specific, Incomplete, and Stubborn Metaphysical Presumptions", *Philosophical Psychology*, **13**, 3, pp. 277–297.

Carey, S. and Spelke, E. 1994, "Domain-Specific Knowledge and Conceptual Change", in L. Hirschfeld and S. Gelman, eds, *Mapping the Mind: Domain Specificity in Cognition and Culture*, Cambridge: Cambridge University Press, pp. 169–200.

Carruthers, P. 2006, *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*, Oxford: Clarendon Press.

Darwin, C. 1881, "Letter 3230 – Charles Darwin to William Graham. July 3rd, 1881", online at: www.darwinproject.ac.uk/entry-13230 [accessed 17 March 2015].

De Cruz, H. and De Smedt, J. 2007, "The Role of Intuitive Ontologies in Scientific Understanding: The Case of Human Evolution", *Biology and Philosophy*, **22**, pp. 351–368.

De Cruz, H. and De Smedt, J. 2012, "Evolved Cognitive Biases and the Epistemic Status of Scientific Beliefs", *Philosophical Studies*, **157**, pp. 411–429.

De Cruz, H., Boudry, M., De Smedt, J. and Blancke, S. 2011, "Evolutionary Approaches to Epistemic Justification", *dialectica*, **65**, 4, pp. 517–535.

Descartes, R. 1641, Meditations on First Philosophy, Translated by J. Veitch, 1901, *The Method, Meditations and Philosophy of Descartes*, New York: Tudor Publishing Company, pp. 205–280.

Fales, E. 1996, "Plantinga's Case Against Naturalistic Epistemology", *Philosophy of Science*, **63**, 3, pp. 432–451.

Fodor, J. 1983, *The Modularity of Mind: An Essay on Faculty Psychology*, Cambridge, MA: MIT Press.

Fodor, J. 1985, "Précis of *The Modularity of Mind*", *Behavioral and Brain Sciences*, **8**, pp. 1–42.

Gigerenzer, G. 1991, "How to Make Cognitive Illusions Disappear: Beyond Heuristics and Biases", in: M. Hewstone and W. Stroebe, eds, *European Review of Social Psychology*, **П**2, Chichester: Wiley, pp. 83–115.

Goldman, A. 1975, "Innate Knowledge", in: S. Stich, ed., *Innate Ideas*, Berkeley: University of California Press, pp. 111–120.

Griffiths, P. and Wilkins, J. in press, "When Do Evolutionary Explanations of Belief Debunk Belief?", in: *Contributions from the John Reilly Center for Science, Technology and Values*, Notre Dame: Notre Dame University Press.

Haselton, M and Nettle, D. 2006, "The Paranoid Optimist: An Integrative Evolutionary Model of Cognitive Biases", *Journal of Personality and Social Psychology Review*, **10**, 1, pp. 47–66.

Hume, D. 1748, *An Enquiry Concerning Human Understanding*, Oxford: Clarendon Press.

Kahneman, D. 2011, *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.

Lakoff, G. and Johnson, M. 1980, *Metaphors We Live By*, Chicago: University of Chicago Press.

Law, S. 2012, "Naturalism, Evolution and True Belief", *Analysis*, **72**, 1, pp. 41–48.

Lipton, P. 2000, "Tracking Track Records", *Proceedings of the Aristotelian Society – Supplementary Volume*, **74**, 1, pp. 179–205.

McKay, R. and Dennett, D. 2009, "The Evolution of Misbelief", *Behavioral and Brain Sciences*, **32**, 6, pp. 493–561.

Mithen, S. 1996, *The Prehistory of the Mind: The Cognitive Origins of Art and Science*, London: Thames and Hudson.

Nozick, R. 1981, *Philosophical Explanations*, Cambridge, MA: Harvard University Press.

Papineau, D. 2000, "The Evolution of Knowledge", in: P. Carruthers and A. Chamberlain, eds, *Evolution and the Human Mind: Modularity, Language, and Metacognition*, Cambridge: Cambridge University Press, pp. 170–206.

Pinker, S. 1997, *How the Mind Works*, New York: W.W. Norton and Company.

Pinker, S. 2007, *The Stuff of Thought: Language as a Window into Human Nature*, New York: Viking/Penguin.

Plantinga, A. 1993, *Warrant and Proper Function*, New York: Oxford University Press.

Potts, R. 1996, *Humanity's Descent: The Consequences of Ecological Instability*, New York: Avon.

Russell, B. 1912, *The Problems of Philosophy*. 1988 ed. Great Books in Philosophy. Buffalo, NY: Prometheus Books.

Spelke, E. 1991, "Physical Knowledge in Infancy: Reflections on Piaget's Theory", in: S. Carey and R. Gelman, eds, *The Epigenesis of Mind*, Hillsdale, NJ: Erlbaum, pp. 133–169.

Sterelny, K. 2003, *Thought in a Hostile World: The Evolution of Human Cognition*, Malden: Blackwell.

Stewart-Williams, S. 2005, "Innate Ideas as a Naturalistic Source of Metaphysical Knowledge", *Biology and Philosophy*, **20**, pp. 791–814.

Tooby, J. and Cosmides, L. 1992, "The Psychological Foundations of Culture", in: J. Barkow, J. Tooby and L. Cosmides, eds, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford: Oxford University Press, pp. 53–115.

Vlerick, M. 2012, "How Can Human Beings Transgress Their Biologically Based Views?", *South African Journal of Philosophy*, **31**, 4, pp. 717–735.

Von Uexküll, J. 1909, *Umwelt und innenwelt der tiere*, Berlin: Springer.

Wilkins, J. and Griffiths, P. 2012, "Evolutionary Debunking Arguments in Three Domains", in: G. Dawes and J. Maclaurin, eds, *A New Science of Religion*, New York: Routledge, pp. 133–146.

Wolpert, L. 1992, *The Unnatural Nature of Science*, Cambridge, MA: Harvard University Press.