Philosophic Exchange

Volume 44
Number 1 *Volume 44 (2014)*

Article 2

2014

Evolutionary Theory and Morality: Why the Science Doesn't Settle the Philosophical Questions

William J. FitzPatrick
University of Rochester, william.fitzpatrick@rochester.edu

Evolutionary Theory and Morality: Why the Science Doesn't Settle the Philosophical Questions[1]

William J. FitzPatrick
University of Rochester

Four decades ago, E.O. Wilson famously declared that "the time has come for ethics to be removed temporarily from the hands of the philosophers and biologicized" (Wilson 1975, 562). Fortunately for those of us gainfully employed as philosophers working in ethics, not everyone listened. Still, one finds Wilson's idea echoed frequently in popular science writing today. Consider this passage from a recent article in *The Economist*, bearing the provocative headline: "Biology Invades a Field Philosophers Thought was Safely Theirs". The author writes:

> Whence morality? That is a question which has troubled philosophers since their subject was invented. Two and a half millennia of debate have, however, failed to produce a satisfactory answer. So now it's time for someone else to have a go...Perhaps [biologists] can eventually do what philosophers have never managed, and explain moral behavior in an intellectually satisfying way.[2]

While I'm not going to deny that evolutionary biology and other sciences have important things to tell us about morality, in some sense, I think there is a lot of confusion about what exactly they can tell us, and how much they can tell us. To begin with, we need

---

[1] This article derives from a lecture designed as a concise and broadly accessible presentation of ideas developed more fully in FitzPatrick (2008, 2011, and 2014a, b), and thus draws significantly from that work.

[2] "Moral Thinking: Biology Invades a Field Philosophers Thought Was Safely Theirs," *The Economist*, February 21, 2008, accessed on August 29, 2008 at: http://www.economist.com/science/displaystory.cfm?story_id=10717915.

to think carefully about what exactly we even mean by 'morality' in these discussions, and recognize the very different questions scientists and philosophers ask about it, often in rather different senses. The above passage makes it sound as if there is a single set of questions that philosophers have raised and failed to answer, which will now thankfully be settled for us by evolutionary biologists. But that is a confusion and it leads to a great deal of overreaching in the name of science, both by some scientists and by some philosophers who appeal to the science to advance certain deflationary philosophical conclusions about morality.

My aim here is first to make some distinctions and sort out some issues, and then to examine one overreaching claim in particular, about the explanation of our moral beliefs by appeal to evolutionary causal influences. That is a claim used by some philosophers to argue that evolutionary biology somehow forces on us either a skeptical or a purely subjectivist understanding of morality. I will try to explain why I think this is misguided and is a poor use of science in philosophy.

1. First of all, then, what is the plausible part of the thought that animates people like Wilson and the author of that passage from *The Economist*? I think it is basically this: once we accept that natural selection has shaped not only the physiologies of organisms but also patterns of social behavior, in animals ranging from ants to chimpanzees, it is natural to suppose that there have been similar Darwinian influences on human nature as well: we should fully expect the human brain and mind to have been molded by social selection pressures to perform in ways that were adaptive in ancestral environments. This then opens the door to rethinking traditional explanations of human cognition, emotion and behavior (Wilson 1978). And in particular, we might wonder whether many phenomena falling under the heading of 'morality' can be freshly illuminated by approaching them from the perspective of evolutionary biology.

Following Wilson's lead, then, evolutionary psychologists view a wide array of psychological traits, including moral feeling and judgment, as culturally conditioned expressions of traits that have a genetic basis and evolved through natural selection. Why, after all, do human beings have a capacity for moral judgment and motivation in the first place, which appears to be no less universal than our capacity for language? Why do we experience emotions like sympathy, parental love, jealousy, resentment and guilt, and

cross-culturally make judgments condemning incest, unfairness, cheating, and in-group harming, or praising loyalty, reliability, and cooperation? The suggestion is that these things stem ultimately from psychological adaptations—that is, dedicated (or 'domain specific') psychological mechanisms that originally emerged as adaptations for solving problems of social exchange faced by our Pleistocene ancestors, though they may or may not remain adaptive in current environments.[3]

Again, I'm not going to deny that there's some truth in that general idea. But the conclusions that actually follow from the scientific work are likely to be much less flashy and more heavily qualified than the triumphalist talk of finally 'explaining morality' scientifically. So let's start by looking more closely at what it even means to speak of 'explaining morality' scientifically, and the rather different concerns philosophers might have about morality in a somewhat different sense.

2. Another recent headline, this time from CNN, is helpful here. It asks: "Can morality be changed magnetically?" (Landau 2010) Now when philosophers hear this they just laugh, not because there is no interesting question here, but because whatever is meant by 'morality' here is obviously not at all the same thing philosophers typically mean when they talk about morality. A philosopher might ask, for example: "Does morality require that we make large sacrifices to aid distant strangers?" or "are there objective truths about morality"—moral truths about how we should live, or about the wrongness of bullying or rape or slavery, for example. Clearly we're talking about very different things.

The CNN story is about the effects of transcranial magnetic stimulation on the actual moral judgments people make. So 'morality' in that headline refers to certain empirical phenomena: namely, people making certain types of moral judgment under certain conditions. Apparently magnets can affect the judgments people make, and we can observe this phenomenon and study it. So the scientists are asking how certain empirical phenomena involving human thought and behavior are affected by certain

---

[3] Barkow, Cosmides, and Tooby (1992); Buss (2005); but for critiques of evolutionary psychology, see also Coyne (2000), Buller (2005), and Richardson (2007).

stimuli. [4] And more generally they ask questions like: which part of the brain is responsible for producing those phenomena, or for short: which part of the brain is responsible for 'morality'? Similarly, when people ask 'whence morality?' as in that *Economist* passage, or speak of the 'evolution of morality', they're asking questions about the causes of certain empirical phenomena. They are interested in explaining how we came to have a capacity for moral judgment or a disposition to make certain types of moral judgment. So in all these cases we're talking about 'morality' in what I'll call the *empirical* sense: morality as a certain set of empirical phenomena having to do with moral judgment, to be causally explained scientifically.

By contrast, moral philosophers are not typically or primarily concerned with causal explanations of morality in the empirical sense. When we speak of morality, whether we're asking first-order questions about what morality requires of us, say, or second-order questions about whether there are objective truths about morality in that sense, we are speaking of 'morality' in what I'll call the *normative* sense. In this sense, 'morality' refers roughly to however it is we ought to live—that is, to a set of norms that *ought* to be adopted and followed, whether they actually are or not.[5] This is

---

[4] The empirical finding was that when a certain part of the right side of the brain is disrupted, people give intentions less weight in moral judgment, focusing more on outcomes.

[5] Although I'm calling this the 'normative' sense of 'morality', not all questions about morality in this sense are normative or first-order ones. The first question above (about what morality requires of us) is a normative one, but the second (about whether there are objective truths about what morality requires of us) is not: it is a *metaethical* or second-order question *about* morality in the *normative* sense, namely whether there are objective truths about what is good and bad, right and wrong. Another metaethical question involves the meaning of moral judgments. For example, when I say that morality requires us to help distant strangers, we might ask: am I expressing a belief with propositional content that can be true or false, or just a non-cognitive mental state? One might also construe this metaethical question as being about morality in the empirical sense: our moralizing is an empirical phenomenon, and we want to know what kind of meaning it has. But our concern here is not in any case with causal explanations of the empirical phenomena, and

not an empirical phenomenon to be causally explained, or something that might be affected by magnets. To speak of morality in the normative sense is to speak of how we really ought to live, and this is what you're doing when you say, as a committed moral agent, that morality prohibits bullying or rape, for example, or more simply that such things are morally wrong or immoral. And philosophers likewise typically speak of morality in the normative sense, both when we're doing first-order moral theory and when we're doing second-order theorizing about the nature of morality.

For example, when a philosopher gives a theory to justify the first-order moral judgment that bullying or rape are wrong, say by appealing to the Kantian idea that persons are ends in themselves, she is explaining why morality in the normative sense has a certain content; she is not causally explaining some empirical phenomenon. And when a philosopher considers whether it is an objective truth that morality prohibits bullying or rape—an objective truth that these things are genuinely morally wrong—say by looking at the meaning of moral judgments and possible grounds of moral truth, she is likewise speaking of morality in the normative sense. She is asking whether there are objective truths about how we ought to live.

3. So we can distinguish between scientific explanatory projects having to do with morality in the empirical sense, and various philosophical investigations of morality in the normative sense. These are both valuable projects, but the question I'm interested in here is what relevance each might have for the other.

The simplest point here is a familiar one I'll mention just to set aside. Scientific explanations of why we evolved to have certain dispositions obviously cannot settle first-order moral questions for us about how we should live. For example, a strong disposition to philander appears to be a biological adaptation in male superb fairy wrens, among other species, and suppose it were established that it is also an adaptation in male human beings, which evolved for the same Darwinian reasons. We might even imagine that we evolved a psychological disposition to adopt a double standard whereby male infidelity is regarded as more acceptable than female infidelity,

---

what we're interested in is specifically the meaning of judgments with a certain content—that is, judgments whose *content* concerns morality in the *normative* sense.

and this may have affected various cultural structures as well. Even if all that were true (and it makes no difference to the present point whether it is or not), it obviously wouldn't show that any of this is actually good or right, or give us good reason to act on that predisposition. Such facts about evolutionary influences on human psychology and behavior wouldn't have any *normative or rational force* for us today as we reflect intelligently on how it is good and right to live.

Evolutionary facts may *explain* a lot, but they don't *justify* behaviors or attitudes, and so they don't answer moral questions for us. They might of course have some relevance for our deliberations: we always have to deliberate in light of the facts of life, and evolutionary theory contributes factual information about human life, some of which may well be morally relevant. But it certainly doesn't settle anything for us about what we should strive for. So biologists shouldn't (and typically don't) purport to be to 'explaining morality' in the sense of explaining the *content* of morality in the *normative* sense: that would amount to *justifying* a certain conception of morality in the *normative* sense, and biology can't do that. Such a thing still falls under the philosopher's job description. What biology may hope to do is to help explain at least some aspects of morality in the *empirical* sense.

Here I want to focus on two main questions about the relation between biological and philosophical projects.

(1) How do scientific explanatory projects bear on metaethics (philosophical thinking about the status of morality)—for example, questions about whether or not there are objective and knowable moral truths; and:

(2) How might philosophical issues bear on scientific explanatory projects, perhaps imposing significant limitations on them?

So let's turn to those questions.

4. It is no surprise that scientists approach morality in the empirical sense, like anything else, as a set of empirical phenomena to be causally explained with scientific tools. And there are two broad explanatory projects we need to distinguish here. The first deals with the emergence in human history of basic mental capacities relevant to making moral judgments. How and why did human beings come to have a capacity for normative (and ultimately moral) guidance? That is, how did we come to be able to grasp and accept and internalize norms of social interaction and response,

making and being motivated by normative 'ought' judgments and ultimately moral judgments, with the help of reactive emotions such as resentment, blame and guilt? Where did all of this come from?

It seems clear enough that this calls for a causal explanation of a scientific sort. Since we are evolved creatures, there is presumably some sort of evolutionary explanation for the emergence of such complex psychological capacities, just as with our linguistic capacies. Philip Kitcher, for example, argues that the capacity for moral guidance or governance is a biological adaptation that evolved because it promoted social coordination: it helped to remedy altruism failures in ancestral humans, making possible cooperative projects in larger groups with more reliable conformity, which in turn enhanced biological fitness for members (Kitcher 2006a,b, 2011. Cf. also Gibbard 1992). This is plausible enough and is in any case something moral philosophers should be happy to leave to biologists to speculate about (keeping in mind that it is indeed quite speculative). Scientists can profitably take up these questions about the evolution of basic cognitive and emotional capacities independently of the concerns of moral philosophers because these really just are scientific questions about the origins of certain human capacities. Scientists can investigate this without input from moral philosophers just as they can investigate the causal origins of the basic capacities that enable us to do math without having to engage in inquiry requiring a consult with mathematicians or philosophers of mathematics.[6]

---

[6] Among other things, scientists might debate whether the capacity in question is an adaptation or a by-product of other, more general adaptations. There is also the closely related question whether the basic capacity for moral judgment is *innate* or not, with capacity (or faculty) *nativists* claiming it is (Dwyer 2008) and *antinativists* denying this (Prinz 2008). This would just be the same question as before if a trait's being an adaptation were essential to its being innate, and some take it that way. But we might instead have a conception of innateness whereby it is sufficient for a trait's being innate that it is a consistent and stable feature of the evolved package such that it will be manifested in normal development in virtually any cultural context—never mind whether it was itself selected for *as such*. In this sense, an underlying competence for morality might be innate,

Things get more complicated, however, when we shift to explaining not just the origins of the capacities we employ in moral judgment, but the causes of the particular moral *judgments* we make, or the *contents* of our moral beliefs as we exercise those capacities. I don't here just mean complications stemming from the fact that our moral judgments or beliefs are conditioned by cultural influences as well. Obviously that's true: no one a hundred years ago made moral judgments about the morality of stem cell research or twerking. Both the subject matter of our judgments and the very moral concepts we employ are often heavily conditioned by our cultural influences. But evolutionary psychologists are well aware of that. Their claim is going to be that despite the many cultural contributions, there are biological adaptations in the form of evolved psychological traits that have deeply influenced the more basic content of our moral feelings, judgments and behaviors. In fact, these evolved psychological adaptations have also shaped the cultural developments themselves, so even those cultural influences already reflect deeper biological shaping.

For example, a sexual aversion toward close kin likely evolved for obvious Darwinian reasons, and this universal psychological adaptation may explain moral sentiments relating to incest by influencing emotionally laden moral judgments about incest, which in turn influence behaviors relating to it. Part of that behavior is the development of cultural norms and practices regulating incest, which will therefore themselves be partly steered by these Darwinian influences, perhaps in combination with other adaptations for "regulat[ing] the sexual behavior of one's kin," all mediated by cultural inputs (Lieberman, 2008). So despite the complications of cultural inputs, evolutionary psychologists will still claim to find many Darwinian influences underlying much of the basic content of our moral judgment and resulting behavior. Maybe they haven't fully appreciated the degree of cultural input and variation, as some social scientists (such as Joseph Henrich) have argued. But my concern is with a different sort of complication.

5. It has to do with the fact that our moral judgments or beliefs are not merely things that are *caused* in us, the way human cognitive

---

as opposed to a contingent cultural construction, even if it is not itself an adaptation.

capacities were simply caused to develop over evolutionary time. When you make a moral judgment or form a belief—for example, that racial or gender discrimination in voting or educational opportunities is wrong—this isn't just something that was caused to happen in you, like a sneeze or a rash. Our first question about it shouldn't be: 'what caused this judgment to occur in you?', as if it were just any other empirical phenomenon. As with other judgments or beliefs, when you make a moral judgment or hold a moral belief, you typically have a *reason* for it—a consideration you take to justify it, helping to show why it's true, rather than just some factor that causally explains its occurrence in you. If I ask you to explain why you believe what you do, you'll give me your reason for holding that belief, which again functions as a justification, a kind of defense of your belief purporting to show why it's true; you're not merely citing something that caused it, like some dust that caused you to sneeze. So if we want to explain your moral belief—why you believe what you do about the wrongness of discrimination—we should start by finding out *your reason*, i.e., the consideration you took to justify your belief by supporting its truth. We at least start with what is called a 'reason-giving explanation', as opposed to a merely causal explanation offered from the outside.

The same is true for other beliefs, such as mathematical ones. If we want to explain why you believe that there are infinitely many prime numbers, for example, we don't start by looking for physical causes or merely sociological influences. We start by asking you your reasons. When we ask: why do you believe that there are infinitely many primes, we mean: what reasons do you have for thinking this mathematical claim to be true? And you'll answer our 'why?' question by giving your reasons, which amount to the considerations you take to justify your belief, supporting its truth. For example, you'll cite Euclid's famous proof, which you take to show that there are infinintely many primes.

Now the point is not that morality is just like math. The point so far is just that when we set out to explain beliefs—why people believe what they do—we cannot just start by casting about for *causes* having nothing to do with the *truth* of the *content* of the beliefs in question. There may be some cases that are like that, where that is all that can be said: 'Why does Oscar believe his hat is on fire? It's because a hypnotist caused him to have that belief.' And I grant that there are some cases of moral belief that are likewise just caused by extraneous factors, where the reasons people give for them are just what Jonathan Haidt (2001) calls "post hoc

rationalizations" rather than genuine explanations. For example, someone might judge that homosexuality is evil, and this may just be an expression of his gut reaction of disgust when he imagines it, projected onto the world, and then he casts about for 'reasons' to rationalize his judgment after the fact. Such a thing is not uncommon. But there is no good reason to think that this is all that's *ever* going on when we make moral judgments, especially reflective ones for which we have plausible rationales that withstand critical scrutiny. We cannot just assume that moral judgments across the board typically fit the merely post hoc rationalization model just because some plausibly do.

In trying to explain someone's moral beliefs, then, we should at least start by looking to the person's reasons for holding the belief. And the point is that this gets us to questions of *justification* and not just *causation*, since the person believes what she does because she takes a certain consideration to be a *good reason* that justifies the belief, by showing it to be true. So we then need to ask whether she is right about that, and how she came to be right about that, or where she went wrong. In the math example, those questions get us into substantive mathematics: we are not simply in the business of appealing to physical or sociological causes to explain beliefs about prime numbers, as when explaining a person's sneeze or accent. We instead need to consider whether the proof a person took to be a good reason for believing the mathematical claim *really was* a good reason or not, and why she *took* it to be. And my point is that the same is true in the case of moral belief—or at least we can't simply assume it is not.

You believe, let's suppose, that racial or gender discrimination in voting laws is wrong, or that bullying is wrong. Why? Well, it could be that a hypnotist arbitrarily caused you to have that belief, and that's all there is to it. But ordinarily that is not how it is. Normally, you have reasons for your moral beliefs. These might include the fact that race or gender are irrelevant to what matters to responsible voting, or the fact that bullying is cruel, causes needless suffering and misery, and is an affront to the dignity of the victim. These reasons are things you take to be *wrong-making* features of discriminatory voting laws or of bullying. So our explanation of your belief goes by way of your judgment that those things are wrong-making features and so *good reasons* for believing the moral claims in question, *justifying* them by showing why they're *true*—why it's true that discrimination or bullying are wrong. We

don't just cite external causes in the way scientists are accustomed to doing for other types of phenomena.

Now why does all this matter? It matters because in philosophy there is a still very live debate about the status of moral judgments, with many—probably most—philosophers maintaining that moral judgments *can* be straightforwardly true or false, just like other claims: there are moral facts alongside other facts. Of course, plenty of people are also skeptical about that, but never mind whether it's ultimately correct or not: it doesn't actually matter for my purposes here. It is enough to recognize that it is at least an open philosophical issue, which means that one cannot simply assume, as a *scientist*, say, that there are no moral truths—that morality is all just a matter of subjective preferences, or expressed attitudes, or mere convention.

Moral nihilism is of course one philosophical possibility, but it is hardly the default, and it is equally possible that there is truth or falsity in morality as in other things. And the point is that if there *is* real truth in the moral domain, then that will likely make a difference to how at least some of our moral judgments or beliefs are properly explained, just as in other domains such as mathematics. So the answer to this philosophical question will make a crucial difference to explanatory projects concerning our moral beliefs, even though our moral beliefs are observable, empirical phenomena. That's the key issue.

To see this, return to math, where there certainly seems to be truth and falsity. We don't look to explain your belief about prime numbers simply by citing evolutionary causes or sociological causal influences, though of course they will be part of the story. We cite primarily *your reasons* for holding the belief, and your judgment that these are indeed *good* reasons for holding it—good reasons to think there are infinitely many primes. And this again isn't merely a judgment you were caused to have by extraneous evolutionary or sociological causes, as by a hypnotist: you came to that judgment because you're mathematically *competent* and correctly *recognized* the proof to be sound, entailing the truth of the proposition in question. In other words, you believe that there are infinitely many primes *because there are*, and being competent, you've grasped that mathematical fact by grasping the *reasons* why there must be infinitely many primes. In other words, the *truth* of the *content* of your belief is relevant to the explanation of it here: it's not that extraneous causes operating independently of the truth of the content of the belief have merely pushed you to have this belief.

The point, then, is that the same thing may hold for your moral beliefs if there are likewise facts about morality, as many philosophers believe there are. Suppose bullying is genuinely, objectively wrong, and is made wrong by the fact that it causes needless suffering and misery, and violates the dignity of the victim, as I've suggested. In other words, suppose it is an objective moral fact that these features of bullying are wrong-making, and that bullying is therefore wrong. And now suppose also that you believe that bullying is wrong. We then ask: why do you believe that? Again, the ordinary explanation here will be a reason-giving explanation. So let's say you give as your reasons precisely these features of bullying, which you correctly take to be wrong-making features, and therefore *good reasons* to believe that bullying is wrong. The next question, then, is why do you believe that?

Well, it could be that you just coincidentally happen to believe those things for reasons having nothing to do with their truth: you believe that causing needless suffering and misery is wrong-making not because it is, but simply because a hypnotist caused you to have that belief, or because evolution or sociological influences coincidentally caused you to believe it regardless of its truth. But the far more natural explanation here—and the one you should find far more plausible if you take yourself to be a competent moral agent—parallels the explanation in the math example, again assuming for the sake of argument that there are truths about these things. You believe that bullying is wrong, I suggest, *because* it *is* wrong and, being morally competent, you've *recognized* that moral fact by grasping the *reasons* why bullying is wrong, correctly recognizing the wrong-makingness of inflicting needless suffering. This, in any case, is the kind of explanation many philosophers will give in cases like this. And if it is correct (which is at least an open possibility), then the proper explanation of at least some of our moral beliefs—where we are getting things right for the right reasons—essentially appeals to our recognition of moral truths. In other words, *the moral truths will come into the explanation of our beliefs.*

Now this is important because it obviously contrasts with scientific explanations of moral beliefs. Scientists don't trade in exotic items like moral truths or processes like someone's grasping a moral truth by recognizing good reasons for holding a moral belief. Those aren't empirical categories. And that's fine, of course: I'm not saying scientists should talk about those things; that's for philosophers. What I'm arguing is that since the philosophical picture I've sketched remains a live, open possibility,

we have to keep that possibility in mind when approaching the question of how best to explain our moral beliefs.

If nihilism were true, contrary to the picture I've just been exploring, then things would be simpler: the only explanations of our moral beliefs would be scientific causal ones. Scientists could go ahead and give purely causal explanations of why we took certain things to be good reasons for holding certain moral beliefs even though they aren't. This would just be a matter of evolutionary and sociological influences, on the very same model as the one we would employ in explaining why a bigot takes takes the fact that a certain marriage is mixed-race to be a good reason to believe that it's wrong. If there is really no such thing as wrong-makingness, then all judgments about it—whether your plausible judgments or the racist's deplorable ones—will just be explained on the same model, by appeal to things like evolved instincts and cultural pushes and pulls this way or that way.

As I have emphasized, however, nihilism is hardly the obvious default here, which means that *scientists cannot just proceed as if it were correct*. The existence and nature of moral truth is a core philosophical issue, and nihilism is just one highly contentious philosophical position. If nihilism turns out to be false and there are moral truths of the sort I've suggested, and we're capable of competently grasping them, then that will plausibly be a crucial part of the best explanation for why we believe at least some of what we do morally speaking. So if someone tries to explain why you believe that bullying is wrong simply by appealing to evolutionary or other causal factors that pushed you incidentally in that direction, with no appeal to your having competently grasped the moral truth that bullying is wrong by understanding the reasons why it is wrong, then that explanation may be fundamentally lacking, just as it would be in the parallel math case. Indeed, if you are not a nihilist and are serious about your belief then you should feel rather insulted by such explanatory claims about why you believe what you do, which appeal simply to extraneous causes. Such explanations are ultimately more condescending than "intellectually satisfying" (as they are claimed to be in the *Economist* passage we started with).

Of course, in the case of *false* moral beliefs, such as the belief that interracial marriage is morally wrong, there will be nothing but merely causal factors at work, since there was no relevant moral truth to have been grasped in those cases. The explanations in *those* cases would look just like all explanations would under nihilistic

assumptions. My point is that we can't just assume that *all* explanations should look like that, across the board—at least not as long as it is an open philosophical possibility that there are moral truths and our grasping some of them has guided some of our moral beliefs. And this has an interesting and perhaps rather surprising implication: if scientists are to avoid begging central philosophical questions, they cannot assume that all empirical phenomena admit of purely scientific causal treatment, as one might have thought. Our having certain moral beliefs is, after all, an observable, empirical phenomenon, but under the philosophical possibility I've been describing, some of our moral beliefs won't be explained fully and adequately by appeal to scientific causal explanations. So if we are to avoid begging philosophical questions, as by just assuming nihilism, then scientific explanatory claims concerning our moral beliefs have to be put quite modestly.

No scientist should claim to have 'explained morality' in the sense of explaining why you hold the moral beliefs you do. They can reasonably claim to have identified some causal factors—many of them evolutionary—that have plausibly influenced many human moral beliefs to some degree. But that is a far cry from revealing the causes of all our moral beliefs or showing that they are all morally blind. And the same goes, incidentally, for religious beliefs, if scientists are to avoid begging theological questions that lie outside the scope of their inquiry.

6. Let me conclude by explaining briefly how I think some philosophers have missed these points and made overreaching explanatory claims in the name of science, which they have then tried to use to support deflationary metaethical views. I'll call these philosophers 'evolutionary debunkers'. Their aim is roughly to argue that there can't be moral truths that are both *knowable* and also *objective*—in the sense of not being just a function of our contingent, subjective desires or attitudes. If basic moral truths are objective, as many philosophers think, then according to the debunkers we could never know them, and so we would be stuck with moral skepticism. So either we can have moral knowledge but morality is all subjective, or we can't have any moral knowledge at all (Street 2006). And this is all supposed to be true because evolution is incompatible with our having justified moral beliefs if moral truths are objective. Now why do they think that? Basically the argument goes like this (drawing here mainly on Street 2006,

though see also Joyce 2006 for related arguments for a different but related conclusion).

Evolutionary biology, they claim, has deeply shaped the content of our moral beliefs across the board. It causally explains why we believe what we do. Cultural influences obviously fill in various details, but these are all variations on deeper, basic themes bequeathed by natural selection, as it shaped human moral feeling and judgment. And as noted before, even the cultural influences themselves reflect deeper biological ones. So just as evolutionary factors plausibly gave us our basic moral aversion to incest, they have had a pervasive influence on all the rest of our moral beliefs too. But then, the argument continues, we have to ask: how would natural selection have molded our moral belief-forming dispositions? And the answer is: not specifically to be reliable trackers of objective moral truths. Natural selection rewards traits that, over time and in the overall environment, have effects that increase the relative frequency in the gene pool of the alleles that code for them, by promoting the reproductive success of carriers of those alleles. And this holds for a belief-forming disposition no less than for any other trait: human belief-forming dispositions were shaped by natural selection to produce whatever beliefs promoted the reproductive success of hunter gatherers in ancestral environments. Now in many cases this meant shaping our dispositions to form *accurate* beliefs about the world: for example, accurate perceptual beliefs helped our ancestors navigate their environment effectively, thus promoting survival, thus promoting reproductive success. But while accurate representation of the world comes into the best evolutionary story for our perceptual belief-forming dispositions, it needn't work that way for certain other sorts of belief.

Suppose a genetic variation arose conferring a propensity to believe in God, and this led to greater reproductive success because believers are more likely to follow social rules even when no fellow human is there to enforce them, leading to greater social stability and consequent fitness advantages. This disposition would be favored by natural selection quite apart from the *truth* of the *content* of the beliefs it produces—that is, quite apart from whether those beliefs *accurately represent* a realm of objective theological facts: all that matters to the selection story is that having the belief conferred an adaptive advantage, which it did quite apart from its truth. But the same thing is true for moral beliefs: if the disposition to form certain moral beliefs led to behaviors that enhanced the

propagation of the alleles underlying that disposition, then that disposition would have been favored by natural selection regardless of whether or not those beliefs accurately represented some realm of objective or independent moral truths or facts. Unlike in the perceptual case, all that mattered in the moral case, like the religious case, was having *whatever* moral beliefs led (regardless of their truth or falsity) to behaviors that maximized the relevant genetic propagation. Fitness-enhancing behavioral dispositions would be rewarded regardless of whether they are truly morally right or wrong in the philosopher's sense: philandering or killing step children might be rewarded (as they are in other species) no less than caring for one's offspring. And in the same way, fitness-enhancing dispositions to form moral beliefs will be rewarded regardless of whether the beliefs are true or not. Even if the beliefs that philandering is good or that killing stepchildren is permissible are false, the disposition to form such beliefs would have been rewarded by natural selection if it promoted the relevant genetic propagation, no less than the disposition to form plausibly true beliefs, like the belief that nurturing one's children is good.

The conclusion the debunkers draw from all of this is that our confidence in the accuracy of our moral beliefs—just as in the case of religious beliefs traceable to evolutionary forces—should therefore be undermined. Or at least that is the case if there are objective moral truths, so that our beliefs would have to reliably represent those truths to count as accurate. If our moral beliefs are thoroughly molded by evolution, and evolution shaped them according to Darwinian principles having nothing to do with objective moral truths as such, then we have no reason to think our moral beliefs accurately track objective moral properties and truths, even if they exist. It would be sheer luck if they did. And once we realize that, we lose our justification for our moral beliefs, at least if moral properties and truths are objective matters. So either moral truths are all reducible to subjective facts about us, or else they're objective but we can't know them because evolution has shaped our moral beliefs completely independently of them, in ways we have no reason to think would track them reliably. That's roughly the evolutionary debunking argument, or the most familiar strand of it, though I'm oversimplifying a bit (see FitzPatrick 2014a and 2014b for more detailed discussion, and also in connection with the response below).

7. Now there are lots of things one might say in response, but my central criticism is actually quite simple, given what we have already seen. The problem is that the argument just begs the question against the philosophical position it is supposed to be refuting. It is supposed to be *showing* us that there are *no knowable, objective* moral truths. And it tries to do that by appealing to science to support the claim that the content of our moral beliefs is pervasively shaped by evolutionary causal factors, which would have steered our beliefs in directions having nothing to do with objective moral truths as such. But the problem is that the science can't possibly establish such a strong explanatory claim about the explanation of our moral beliefs. (And it can't do that with regard to theological beliefs either.) The most that evolutionary biology can do is to point to some plausible causal influences in evolutionary history on some of our moral beliefs to some extent. It can't thereby rule out the philosophical possibility I've described that many of our moral beliefs are instead the result of *apprehending moral truths*, on the model of mathematical beliefs arrived at through understanding relevant reasons as such. So the debunking argument never really gets off the ground.

In other words, the debunkers can't just start out assuming there are no knowable, objective moral truths in an argument intended to show that there aren't any: that would be obviously question-begging. But if there *are* knowable, objective moral truths then it is entirely possible and plausible that many of our moral beliefs are properly explained by the fact that we have apprehended those truths by understanding the reasons why they're true, as such. For example, I think that is the proper way to explain your belief that Talibanic treatment of girls and women is unjust and cruel and therefore wrong. And if that is the case, then it is just false that our moral beliefs across the board are simply a result of *morally blind causal shaping* by evolutionary or other cultural forces indifferent to moral facts as such. We can therefore undermine the debunking argument from the start by rejecting the first premise, which overreaches in its strong explanatory claims about 'our moral beliefs' across the board.

The debunkers may be right that insofar as *some* of our moral beliefs reflect nothing but morally blind causal shaping by evolutionary or other forces of the sort scientists can study, then *those* beliefs are unreliable. Evolution would not have shaped us to be generally reliable in that way. But that simply doesn't matter, because we needn't accept that all our moral beliefs are thoroughly

shaped by evolutionary and other blind causal factors. It is entirely possible that in addition to shaping some of our moral beliefs to some degree, evolution also gave us the basic raw materials—reflective, intellectual, and emotional potentialities—necessary for us to develop reliable moral belief-forming dispositions ourselves. And by developing and deploying those dispositions intelligently, we may well have come to have plenty of reliable moral beliefs.

We don't *need* evolution to have *made* us reliable in tracking objective moral truths any more than we need it to have made us reliable in tracking truths about non-linear algebra or quantum non-locality or philosophy for that matter, none of which played any more role in the evolution of Pleistocene human cognitive traits than moral truths did. It is enough if we have been able to develop the basic potentialities evolution gave us through the right forms of experience, training and reflection in rich cultural contexts, in such a way as to come reliably to track moral truths through gaining moral understanding.

8. Let me emphasize in closing that I don't for a moment claim to have shown that there *are* in fact objective and knowable moral truths (though I believe there are). That is a much larger and more difficult question. My purpose here has been relatively modest: namely, to emphasize the live philosophical possibility that there is such a thing as getting things objectively right or wrong in morality as in other domains (though of course with plenty of room for legitimate variation too), and that this possibility of objective moral truth reveals potential *limitations* on certain scientific explanatory projects. It doesn't affect scientific explanations of how we evolved the basic cognitive capacities we use in moral judgment, but it does bear on attempts to explain why we believe what we do when it comes to the substance of morality: that is, it affects explanations of the content of our moral beliefs. We cannot just assume that this can all be done by appeal to the kinds of causal factors scientists work with, such as evolutionary influences. And that means that, contrary to that passage from the *Economist*, it is a mistake to think that philosophy needs rescuing by biology in order to "explain morality in an intellectually satisfying way." Both philosophy and the sciences have valuable things to tell us about morality, but they are quite different things. And what is "intellectually satisfying" or not when it comes to explanations of morality depends partly on the outcome of central ongoing

philosophical debates—debates that will not be settled anytime soon, and not by the sciences in any case.

Bibliography

Barkow, J.H., Cosmides, L. and Tooby, J. (1992), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.

Buller, D.J. (2005), *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature*. Cambridge, MA: MIT Press.

Buss, D.M. ed., (2005), *Handbook of Evolutionary Psychology*. Hoboken, NJ: John Wiley & Sons.

Coyne, J.A. (2000), "Of Vice and Men," *The New Republic* 222(14): 27-34.

Dwyer, S. (2008), 'How Not to Argue that Morality isn't Innate: Comments on Prinz', in W. Sinnott-Armstrong ed., (2008), pp. 407-18.

FitzPatrick, W.J. (2008), "Morality and Evolutionary Biology", in E.N. Zalta ed., *Stanford Encyclopedia of Philosophy*: http://plato.stanford.edu/entries/morality-biology/.

FitzPatrick, W.J. (2011), "Biology, Evolution and Ethics," in C. Miller ed., *The Continuum Companion to Ethics*. London: Continuum.

FitzPatrick, W.J. (2014a), "Debunking Evolutionary Debunking of Ethical Realism." *Philosophical Studies*, January 2014 Preprint. DOI: 10.1007/s11098-014-0295-y.

FitzPatrick, W.J. (2014b), "Why There is No Darwinian Dilemma For Ethical Realism." In M. Bergmann and P. Kain eds., *Challenges to Moral and Religious Belief: Disagreement and Evolution*. Oxford: Oxford University Press.

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

Haidt, J. (2001), "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment', *Psychological Review* 108(4): 814-34.

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

Kitcher, P. (2006a), "Biology and Ethics," in D. Copp ed., *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, pp. 163-85.

Kitcher, P. (2006b), "Between Fragile Altruism and Morality: Evolution and the Emergence of Normative Guidance," in G. Boniolo and G. DeAnna eds., *Evolutionary Ethics and Contemporary Biology*. Cambridge: Cambridge University Press, pp. 159-77.

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

Landau, E. (2010), 'Can Morality be Changed Magnetically?', CNN.com, March 30, http://pagingdrgupta.blogs.cnn.com/2010/03/30/can-morality-be-changed-magnetically/

Lieberman, D. (2008), "Moral Sentiments Relating to Incest: Discerning Adaptations from By-Products," in W. Sinnott-Armstrong ed., (2008): pp. 165-90.

Prinz, J. (2008), "Is Morality Innate?" in W. Sinnott-Armstrong ed., (2008): pp. 367-406.

Richardson, R.C. (2007), *Evolutionary Psychology as Maladapted Psychology*. Cambridge, MA: MIT Press.

Sinnott-Armstrong, W. ed. (2008), *Moral Psychology, Vol. 1: The Evolution of Morality: Adaptations and Innateness*. Cambridge, MA: MIT Press.

Street, S. (2006), "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127: 109-66.

Wilson, E.O. (1975), *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard University Press.

Wilson, E.O. (1978), *On Human Nature*. Cambridge, MA: Harvard University Press.