Human morality: From an empirical puzzle to a metaethical puzzle
Richard Joyce

1. Human moral thinking: An empirical puzzle

It is a striking feature of human psychology that we assess the world in moral terms. We do not merely like and dislike things—we judge certain actions to be *obligatory*, certain people to be *wicked*, certain states of affairs to be *deserved*, and so on. This is a phenomenon that cries out for explanation, for it seems no great feat of imagination to envisage an intelligent social species that does nothing of the sort. Picture a community of beings who are inclined to cooperate with each other (albeit conditionally) because they really like each other, who positively enjoy laboring in service of their fellows' welfare, in whom natural selection has simply expunged any temptation to defect. These beings may not even have the cognitive capacity to grasp our basic moral concepts; maybe we couldn't explain *desert* to them if we tried. Such creatures would be quite different from humans, obviously, though it is difficult to say with confidence *how* different. Our sense of fellow feeling is fainter than theirs, our prosocial resolve far more susceptible to undermining temptations, and, of course, we have the capacity which they lack: to wield the peculiar suite of conceptual tools necessary for moral judgment.

What processes have brought it about that we humans are as we are and not like these imaginary creatures? What explains our unusual tendency to evaluate our world, each other, and ourselves in moral terms? The answer "natural selection" is almost certainly correct, but it's a coarse-grained answer that obscures much of significance. Natural selection is the process responsible for giving us a brain capable of moral thinking (compared, say, to a wombat's brain), but that is not to say that the capacity for moral thinking is an adaptation. It is possible that, rather, human moral thinking emerged as a byproduct of psychological capacities that evolved for other purposes. By analogy, if one were wondering why humans can drive cars but wombats cannot, then "natural selection" is in some sense the right answer, but it doesn't follow that the capacity to drive cars was selected for somewhere in the hominin lineage. Rather, our ability to drive cars supervenes on various traits (largely lacked by wombats) that evolved for other fitness-pertaining reasons—such as visual perception, hand-eye coordination, cognitive mapping, etc.

Distinguishing a trait that is an adaptation from a trait that is a byproduct of adaptations is a difficult business, for no amount of careful examination of the intrinsic features of an organism can settle the matter. Consider Stephen Jay Gould's architectural analogy of *spandrels*: triangular regions of a church's wall that are the inevitable outcome of mounting a dome on a ring of arches (Gould & Lewontin

1979); the spandrels are not a design feature but a byproduct. However, we can imagine an eccentric architect inspired by a vision of beautiful spandrels; in the resultant building it is the arches and the dome that are the byproducts. Picture the two churches side by side: one for which the spandrels are the byproduct, one for which they are the primary design feature. The two churches could be brick-for-brick identical; determining which features are byproducts requires discerning something of the architect's intentions. In the case of Darwinian evolution, determining whether a trait is an adaptation or byproduct requires discerning something of the selective forces that were operational during the period of the trait's emergence. And the problem is, of course, that in the latter case it is usually more reasonable to speak of "educated guessing" than "discerning."

Things would stand differently if we had evidence that the trait emerged relatively recently—too recently for natural selection to operate—for then we could rule out the possibility of its being a genetic adaptation. But we have no such evidence in the case of human moral thinking. As far back as the historical record stretches we find moral judgments. The Egyptian *Book of the Dead* includes the words spoken by the deceased to Osiris: "I have not sinned against men. … I have not wrought evil" (Wallis Budge 1901: 360). The sixth tablet of the *Epic of Gilgamesh* recounts Ishtar describing actions in terms that are translated variously as "evil," "despicable," "foul," or "abominable." There is no evidence to suggest that we should not project human moral thinking back far earlier into prehistory than these first-recorded instances.

Thus we are really in no position to explain moral thinking with any confidence. Note that I am not referring to the explanation of particular moral judgments or moral systems. One might, for example, be able to explain why the ancient Egyptians thought sibling incest morally permissible (perhaps even in some circumstances obligatory) whereas modern Western cultures tend to find it morally repugnant. As to the question "Why does Ptolemy judge incest to be morally permissible?" a reasonable answer may well be "Because he was taught this." I, however, am scrutinizing the matter through a wider lens than this, one focused on explaining the trait that these different cultures and individuals share: Both we and the ancient Egyptians (and every other known culture) employ a conceptual framework that categorizes actions in terms of (inter alia) permission, obligation, and repugnance. Particular moral systems are no doubt shaped by various historical and environmental factors and learned by individuals; the question is whether there is an evolved psychological capacity dedicated to this kind of learning. (After all, wombats can't learn moral systems.) To the extent that we do not know whether this more general trait is an adaptation or a byproduct, our ability to explain human moral thinking remains fundamentally incomplete.

The difficulties inherent in distinguishing adaptations from byproducts have not prevented plenty of plausible speculation on the genealogy of human moral thinking. Some, including myself, have advocated the nativist hypothesis that mechanisms dedicated to moral evaluation emerged as psychological adaptations in the human lineage. (See Ruse 1986; Dwyer 2006; Joyce 2006a; Mikhail 2011.) Others have

argued that moral thinking is a byproduct of psychological mechanisms that evolved for other purposes. (See Nichols 2005; Prinz 2007, 2008, 2009; Ayala 2010; Machery & Mallon 2010.) It is not my intention on this occasion to attempt to adjudicate this complex and tricky debate; I'm inclined to think that the evidence necessary to do so remains unavailable. Rather, what I shall do is make some very broad comments about what both sides of the debate seem to have in common, and then on this basis move the discussion into the province of metaethics.

2. Genealogical hypotheses sketched

What advantage did moral thinking provide our ancestors? If moral thinking is a genetic adaptation, then it is the product either of individual selection or group selection (assuming the two are distinct). If the former, then moral thinking provided its bearers with some fitness-enhancing benefit (relative to the competition). If the latter, then moral thinking provided its bearers' relevant group with some fitness-enhancing benefit (relative to the competition).

One can apparently find the latter view expressed by Darwin:

Although a high standard of morality gives but a slight or no advantage to each individual man and his children over other men in the same tribe, … an advancement in the standard of morality will certainly give an immense advantage to one tribe over another. (Darwin [1879] 2004: #)

I say "apparently" because Darwin's mention of "a high state of morality" is vague (for our purposes); it is unclear at this point in his discussion whether he is targeting the trait of moral *judgment* in particular.[1] This is often what one finds in those who advocate a group-selectionist account of the evolution of "human morality" (e.g., Gintis et al. 2008; Krebs 2011; Wilson 2012): they turn out to be attempting to explain human *cooperation* or human *altruism*—neither of which necessitates the capacity to evaluate the world in distinctively moral terms such as *obligation*, *wickedness*, *desert*, and so on. (Remember that the amoral creatures described earlier were cooperative and altruistic.) Nevertheless, advocates of such views may well assent to the natural thought that moral thinking operates largely in the service of cooperation. In other words, if certain forms of human cooperative behavior emerged via a process of group selection, then plausibly moral thinking emerged as a proximate mechanism for governing that behavior. This provides an immediate rough answer to the question posed at the start of this section: The advantage that moral thinking provided our ancestors was that it encouraged certain group-adaptive forms of cooperation.

A similar answer is also widely endorsed by those who eschew group selection and seek to explain the evolution of morality in terms of individual selection (see Ruse 1986; Frank 1988; Irons 1996; Joyce 2006*a*, 2006*b*; Churchland 2011): Here again it is widely accepted that the adaptive function of moral thinking was its contribution to social cohesion. Individual-selectionists face the additional challenge

---

[1] It is also unclear whether Darwin thinks the human moral sense is an adaptation or a byproduct.

of explaining how cooperation produces benefits that accrue *to the individual*—in particular, they face the problem of why a disposition to defect would not better serve the individual—but evolutionary biology has made great strides in resolving this challenge in recent decades. Kin selection, reciprocity, and mutualism (inter alia) are processes by which individual-level selection can favor forms of cooperation, and therefore to the extent that moral thinking can contribute to these forms of cooperative behavior, we have a broad answer to how moral thinking might have been advantageous to our ancestors on an individualistic basis.

There remains an important puzzle as to *how* moral thinking contributed to cooperative behavior. The puzzle can be framed by asking why humans didn't evolve simply to *want to* cooperate (albeit conditionally). One can answer by saying that humans *did* evolve to want to cooperate (albeit conditionally), and moral thinking is one special psychological mechanism dedicated to regulating this desire. Conceiving of an action as one that "must be done whether I like it or not"—as opposed to categorizing it simply as furthering one's desires—is likely to strengthen one's resolve to perform it; the former is less prone to self-sabotaging temptations. Moral thinking has a distinctive emotional profile: Failure to perform an action that one simply wants to do leads to regret; failure to perform an action that one judges to be morally obligatory leads to *guilt*—and guilt encourages motivation to restore social equilibrium (Tangney & Fischer 1995). Moreover, observing another person perform an action that one simply wishes s/he hadn't done (e.g., having won a race that you wanted to win) prompts different emotions than observing someone perform an action that one judges to be morally wrong (e.g., having won the race by cheating). The latter licenses not only a different emotional reaction but a different social response—a social response aimed at re-establishing and maintaining social order. Because of this essential social aspect of moral thinking, even an action that one might be inclined to think of as entirely self-regarding can contribute to social coordination. If a community were to judge wearing a type of footwear, say, to be not merely foolish or unfashionable but *morally wrong*, then members of the community would feel warranted in taking an interest in any footwear-transgressors. Abiding by the moral footwear norm could thus be used as a marker of solidarity and social inclusion—as a form of cooperation. For empirical evidence that moral judgment strengthens cooperative tendencies, see Bandura 1999; Tenbrunsel & Smith-Crowe 2008; van der Linden 2011; Simpson et al. 2013; and Bartels et al. 2016.

Even those who take moral thinking to be a byproduct rather than a dedicated adaptive mechanism tend to agree with this general picture of its operating in the service of cooperation. Francisco Ayala (2010), for example, argues that morality emerged not as an adaptation but as an exaptation from traits like the human abilities to anticipate consequences, to make value judgments, and to choose between alternative courses of action. Yet when speculating about why moral thinking would thus emerge, his explanation is in terms of "patterns of action beneficial to the tribe or social group" (2010: 9019) and "promoting social stability and success" (2010: 9021). Jesse Prinz (2007, 2008, 2009) also argues for the non-nativist view of moral thinking as a byproduct of psychological traits like the capacities for higher-order emotions

(i.e., emotional responses to emotions), perspective-taking, and abstraction. Yet when speculating about why moral thinking would thus emerge, his explanation is in terms of avoiding "a potential collapse in social stability" (2008: #); and when considering why some moral systems flourish over others, his explanation is in terms of certain judgments being "essential to forming successful bands and tribes" while others fail to encourage "the level of cohesion needed for survival" (2009: 178).

So those who take moral thinking to be a byproduct do tend to see morality as having a *purpose*—but they see that purpose in terms of serving social ends rather than as an evolutionary function. This important difference aside, though, they are generally broadly in agreement with the adaptationists about why moral thinking emerged: because it helped strengthen social cohesion in ancestral communities. (Byproduct theorists need not say this; they may hold that morality is an accidental byproduct of adaptations that has no particular function at all. Such a view would not undermine any of the arguments that I will develop in the remainder of this chapter.)

I have been at pains to highlight this observation because, in my opinion, it can be pressed into the service of a metaethical conclusion. From my metaethical point of view, what all these sketched genealogical hypotheses strikingly have in common is that they can be happily endorsed by someone who is utterly skeptical of moral truths. When it comes to conducting debates about the origins of moral thinking—about whether it is an adaption or a byproduct, the result of individual selection or group selection, a genetic or cultural trait, and so on—whether one believes that moral judgments are sometimes true or believes them to be systematically false makes not an iota of difference. The essential feature of the genealogical hypotheses on offer is that our ancestors' moral beliefs were, in some sense, *useful*—the substantive debate concerning in what ways they were useful. But it is no essential feature of any of these hypotheses that any of our ancestors' moral judgments were *true*. This forms the basis of an epistemological debunking argument.

3. Human moral thinking: A metaethical puzzle

Let's bring some metaethical actors onto the stage.

First we have the moral error theorist, who takes an attitude toward morality analogous to the atheist's attitude toward religion. The atheist thinks (A) that people engaged in religious discourse generally purport to state truths and express their beliefs, but (B) the world simply isn't furnished with the objects, properties, and relations necessary to render these assertions and beliefs true. Similarly, the moral error theorist thinks that what it would take for a claim like "Stealing is morally wrong" to be true is for stealing to instantiate the property of moral wrongness, but there simply is no such property, so the statement is false. (It doesn't follow, of course, that the error theorist thinks that it's fine to steal. She also thinks that "Stealing is morally good" and "Stealing is morally permissible" are false. And she may be vehemently opposed to stealing on various *non-moral* grounds.)

Next we have the moral realist, who agrees with the error theorist that in order for the judgment "Stealing is morally wrong" to be true, stealing must instantiate the

property of wrongness, but the realist disagrees with the error theorist by maintaining that there is such a property (and the realist probably thinks that stealing usually instantiates this property). In addition, the realist holds a view about the nature of this property: that it is, in some fashion to be specified, a *mind independent* property: The wrongness of stealing is a fact we discover, not a fact that we create in taking a certain attitude toward it.

Lastly, we have the moral skeptic, who holds that there is no such thing as moral knowledge. Let's accept for the sake of argument (bracketing off some well-known complications) that for S to know that *p* is for S to have a true, justified belief that *p*. This tripartite account of knowledge creates space for three kinds of moral skeptic. First, there are noncognitivists, who deny that moral judgments express beliefs. Second, there are those who accept that moral judgments express beliefs but deny that these beliefs are ever true. This is the error theoretic view which we have just encountered (i.e., error theorists are a kind of skeptic). Third, there are those who accept that moral judgments express beliefs but deny that these beliefs are ever justified.

The debunking argument in which I am interested favors the third kind of moral skeptic. The argument is (roughly) that the genealogical hypotheses sketched undermine the justificatory status of moral judgments. Two things must be noted immediately about this argument.

First, although the error theorist has a role to play in the argument (as we shall see shortly), the conclusion of the debunking argument is not the error theoretic view that no moral judgments are true. The conclusion of this debunking argument is compatible with the truth of moral judgments.

Second, since the claim that a belief is unjustified is compatible with that belief's being mind-independently true, the conclusion of the debunking argument is compatible with moral realism. (Analogy: Someone suffering from paranoid delusions may believe that Dr X is out to get him; but if he has no evidence of this then his belief lacks justification; nevertheless it may be true, and objectively so, that Dr X *is* out to get him.) No fewer than six of the chapters in this volume discuss debunking arguments which target moral realism or moral objectivity, and while there do exist such arguments (e.g., Ruse 1986; Street 2006), the one I am focused on is entirely different. Moral *objectivity* plays no role in this argument; it aims to undermine morality *simpliciter*. With these preliminaries sorted out, let's proceed to the epistemological debunking argument.

I said that what is striking about extant evolutionary hypotheses about the origin of moral thinking is that they can be happily endorsed by the error theorist. It is important to see that this observation cannot be generalized. Suppose it turns out that humans have a mechanism dedicated to distinguishing faces from other visual stimuli. (Whether this hypothesis genuinely withstands empirical scrutiny doesn't matter for my purposes; for discussion, see Slater & Quinn 2001.) One way of expressing this hypothesis—though, admittedly, a rather clumsy way—is to say that human innately

have the concept *face*.[2] An advocate of this view owes us an account of why it would have been adaptive for this mechanism to develop in the evolution of the human brain. Such an account would likely mention the importance of social bonding and emotional communication in the very early period of infancy, the stability of the presence of human faces (and their anatomical structure) in infancy, and so on. Whatever the details, any reasonable account of why it was adaptive for our ancestors to have the concept *face* is surely going to presuppose that faces actually existed; it is useful for infants to have a face-seeking mechanism only because this mechanism is likely to put them in perceptual contact with faces. The same thing will go (*mutatis mutandis*) for evolutionary hypotheses about the human ability to distinguish inanimate from living things (see Simion et al. 2008), to assume object continuity (see Baillargeon 2008), to spot specific dangerous animals such as spiders (see Rakison & Derringer 2008), and so on. The error theorist about the phenomenon in question (though in many cases such a view would be utterly bizarre) would be barred from endorsing the adaptational hypothesis. Were we to discover that spiders never existed in the ancestral environment—were we to uncover astonishing evidence that spiders evolved in the last few thousand years—then the hypothesis that it was adaptive for our ancestors to develop an innate spider sensitivity would collapse. The evolutionary hypotheses pertaining to moral thinking are different: they explain the origin of moral thinking in a manner amenable to the moral error theorist.

The upshot of this, I maintain, is not that the error theorist is proven correct, but that the epistemological moral skeptic is bolstered. If one believes that x is P, but the correct explanation for one's coming to have this belief is consistent with nothing ever having the property P (i.e., the correct explanation is something the error theorist about P discourse can endorse), then one's belief lacks justification. This is not to say that this lack of justification must be a permanent state of affairs. Beliefs that are unjustified can gain justification (via, say, the gathering of further evidence). In this case, what the non-skeptic needs to do is show that although the error theorist might *think* that she can endorse the evolutionary hypothesis, in fact she is mistaken. But showing this takes argumentation. In advance of a persuasive case being offered, it seems reasonable to grant the error theorist the benefit of the doubt: If it appears that she can endorse a hypothesis, then we should assume that she can endorse it until it is demonstrated otherwise.

Notice that saying this doesn't amount to granting the error theorist any benefit of the doubt *simpliciter*; rather, it is just a matter of accepting what the moral error theorist can and cannot endorse. Can the moral error theorist endorse the proposition that Napoleon lost the Battle of Waterloo? It would certainly seem so. I suppose someone might deny this, but then at least we should demand to see this person's reasoning, and in advance of being persuaded we should stick to the view that the error theorist can endorse the proposition. (Obviously, allowing this in no way implies

---

[2] I think it is equally clumsy for the moral adaptationist to express the hypothesis by saying that humans innately have the concepts *moral obligatoriness*, *desert*, *moral wrongness*, etc. But for some purposes such expressions can be taken as a kind of shorthand for a much more sophisticated hypothesis.

any sympathy toward the error theoretic view. Allowing that people who believe in ghosts or phrenology can endorse that Napoleon lost Waterloo in no way implies that one is anything other than 100 percent confidently opposed to these ideas.) By comparison, can the moral error theorist endorse the proposition that Napoleon was morally virtuous? Presumably not. So our question is: Can the moral error theorist endorse the proposition that (for example) human moral thinking evolved as a psychological adaptation that strengthened individuals' commitments to reciprocal exchanges? On the face of it, this looks much more like the question of whether she can endorse that Napoleon lost Waterloo than the question of whether she can endorse that Napoleon was virtuous. To the extent that the moral error theorist owes us an account of why humans everywhere employ a massively mistaken conceptual scheme, the evolutionary hypothesis looks like exactly the kind of explanatory answer she will find attractive.

On the one hand, then, the strength of this debunking argument is ultimately quite modest. If we grant the error theorist the benefit of the doubt that she can endorse the evolutionary hypothesis about the origin of human thinking, then we are provisionally granting that our moral judgments lack epistemological justification. On the other hand, this conclusion is the morality-oriented version of a skepticism no less radical than that espoused by the classical tradition. Ancient skeptics such as Pyrrho, Aenesidemus, and Sextus Empiricus all claimed that we lack knowledge, but it is less clear whether they had any argument to support the stronger thesis that knowledge is unattainable. Sometimes, of course, they got carried away by their own rhetoric. The existence of disagreement was central to the classical skeptic's case, and Sextus (our most complete source on Greek and Roman skepticism) discusses this disagreement using phrases that are translated as "unresolvable impasse" (Mates 1996) and "undecidable dissension" (Annas & Barnes 2000). But no argument deployed by Sextus or any other classical skeptic is actually sufficient to establish this stronger claim. Gathering evidence that disagreement is ubiquitous is one thing; gathering evidence that disagreement is *inevitable* is something else. Correspondingly, showing that our moral judgments lack justification is one thing; showing that they are *unjustifiable* is something else.

4. A nod to Sextus Empiricus

To interpret skeptical arguments as attempts to shift a burden of proof onto the non-skeptic is to recognize their provisional nature. The skeptic can never rest easy, for she never knows what new putatively justification-establishing argument the non-skeptic may come up with. Such non-skeptical cases will have to be dealt with as they arise. It is not surprising, then, that Sextus combines his general case for skepticism with specific attacks on various schools that make claims of knowledge: "Against the Grammarians," "Against the Logicians," and so on. In the same way, the proponent of the epistemological debunking argument, even if she has succeeded in shifting the burden of proof, must be willing to confront those opponents who believe that justification for moral judgments can be supplied. I will close by taking a page out of

Sextus' book and saying a few words specifically against some leading non-skeptical views of moral epistemology. Of course, there is a great deal that could be (and has been) said in general terms against any of these views; I will confine myself to briefly noting for each the potentially undermining impact of genealogical considerations.

Against the Reliabilists: Reliabilists maintain that a belief is justified if and only if it is has a certain kind of history, such as being the product of a process that reliably produces true beliefs. Evolutionary genealogical hypotheses of moral thinking provide the big-picture history of moral concepts, and do so in a way that (it would appear) the error theorist can endorse. If beliefs about X are the product of a process which the error theorist about X can endorse, then they are not the product of a reliable process. The reliabilist who wishes to maintain that moral beliefs are justified therefore has work to do. He must demonstrate that, contrary to appearances, the error theorist in fact cannot really endorse the genealogical hypothesis.

Against the Coherentists: Coherentists maintain that a belief is justified if and only if it coheres in a certain way with other beliefs. A person's moral beliefs may cohere quite well together, giving the appearance of justification. However, beliefs about where our moral judgments come from are themselves items that enter into the coherentist mix. Suppose evidence were forthcoming that human moral thinking is due to a genealogical process that definitely presupposes that many of our moral judgments are true (like, say, the processes that give rise to beliefs about faces or spiders). This, I take it, would strengthen the coherence relations that moral beliefs (generally speaking) would have within the doxastic web. However, this is clearly not the evidential situation with which we are actually faced. Rather, the evidence suggests that moral thinking is due to a process that the error theorist appears able to endorse.

Consider an analogy: A parent judges that his daughter's performance at the musical recital is definitely and objectively the best. This belief coheres well with many of his other beliefs: that his daughter was also the best at last year's recital, that his daughter is exceptionally talented in other ways, etc. But then a new belief enters the mix: the parent comes to appreciate that he judges his daughter's musical accomplishments so positively only because of his extreme partiality toward her; perhaps he is provided with some kind of empirical evidence that he would carry on judging her in this way even if her musical talents were meager. It would seem that it's now incumbent on the parent to form yet another belief: a meta-belief concerning what the relation is between his judgment that his daughter is the best and his realization that he thinks this only because of his own partiality. (After all, just leaving the two beliefs sitting there without forming an opinion about the relation between them seems a step away from coherence.) He has a choice. On the one hand he could think "No, even though I judge her to be best only because I love her so much, she really definitely is the best"; or on the other hand he could think "Hmm, I suppose the fact that I judge her to be best only because I love her so much places a question mark over whether she really is the best." The thing about coherentism is

that either option could, with appropriate adjustments to beliefs elsewhere, figure in a total coherent doxastic package. I suggest, however, that the latter option appears to require fewer adjustments elsewhere. Of course, the parent might prefer the former option and then seek out further sources of evidence as to how good his daughter's ability is. Perhaps she really is the best (even though he thinks this only because of his partiality); perhaps the other parents at the recital will agree to this. The point is, though, that in order to maintain the former option within a coherent doxastic package, the parent has extra work to do.

Against the Intuitionists: The coherentist's traditional rival is the foundationalist, who holds that some beliefs can be non-inferentially justified. If one supposes moral beliefs to be justified on foundationalist grounds, then one holds either that some moral beliefs can be non-inferentially justified, or that moral beliefs can be inferred from non-moral beliefs that can be non-inferentially justified. There are very well-known problems for the latter view; I'll put it aside and focus on the former view—which I shall refer to here as "moral intuitionism." Many moral intuitionists (e.g., Ross 1930; Audi 2004) hold that certain moral propositions are non-inferentially justified in virtue of being (in some sense) self-evident, and it might seem tempting to suppose that there are many self-evident moral beliefs. A lot of moral beliefs seem so obviously true that we are aghast at the prospect of their being denied. ("You don't think torturing babies is *morally wrong*?!" exclaims the error theorist's opponent in a tone of disbelieving outrage.) Some have argued that our confidence in certain moral judgments is so high that we can dismiss any argument that purports to deny them, even in advance of examining the argument's details, so sure are we that one of its premises or inferential steps must be mistaken. (See Enoch 2011.)

But genealogical considerations can upset the moral foundationalist's case. Let me make the point using a quick and slightly dirty analogy. Suppose a person is so extremely confident in the proposition that x is P that he is inclined to classify it as self-evident and requiring no inferential justification. Imagine, though, that the only reason he confidently has this belief is that he was successfully hypnotized to hold it. Upon discovering this fact about his own belief-formation process (and supposing the spell of hypnosis now is broken to the point that he can reflect on the epistemic status of the belief) surely the person should at least doubt whether the proposition that x is P has the status of a self-evident truth.

The general problem here is that the very feature that indicates (according to the intuitionist) a belief's self-evident status—such as our according it a very high level of confidence—may be itself subject to a plausible genealogical explanation consistent with an error theory. For example, an evolutionary explanation for moral thinking is not simply going to account for why we have moral beliefs, but probably also will, when properly fleshed out, account for why our moral beliefs are entrenched and enjoy a high level of confidence—why, indeed, they seem self-evident. But if evidence is available that the feature that inclines us to classify a belief as self-evident is itself the product of a process that appears consistent with the falsity of the belief, then that feature can no longer be appealed to as a ground for self-evidence.

Against the Conservatists: Coherentists and intuitionists tend to favor epistemic conservatism: the view that the mere fact that a belief is held accords it a kind of prima facie justification. Conservatism is reasonable if one can be confident that the beliefs in question are likely to be true. If I know that humans tend to be pretty good at judging the relative size of nearby objects, and I know that Fred believes that a certain nearby object is larger than another, then it is reasonable for me to accord Fred's belief some prima facie justification just in virtue of his holding it. On the other hand, if I know that humans tend to be pretty bad at judging probabilities when calculating risks, and I know that Sally (who is calculating risks) judges a certain outcome to have a high probability, then it is not reasonable for me to accord Sally's belief any justification in virtue of her holding it. Even if I know that Sally's belief is extremely widely shared in the epistemic community, if I have reasonable grounds for thinking that humans tend to get things wrong in this domain, then the mere fact that the belief is widely held provides it with no positive justification.

In the moral case, we tend to think that our judgments are pretty good—more like Fred's belief than Sally's. The prospect that our basic moral beliefs might be false—not just our own beliefs but those shared (often emphatically so) by our peers and loved ones—is one so troubling that we tend to shy away from even confronting it. While a person may acknowledge that she has learned many of her moral norms from her community, this is likely to be accompanied by the inchoate thought that *surely someone somewhere* has *somehow* had access to the moral facts. If this inchoate thought were reasonable, then conservatism might provide moral beliefs with a degree of prima facie justification. The genealogical hypotheses discussed earlier, however, starkly reveal what is not reasonable about this thought, for they provide plausible explanations of how it has come about that we are all emphatically confident about moral beliefs in a manner that appears consistent with a moral error theory. One may, of course, embark on the project of showing that, contrary to appearances, the error theorist cannot endorse the genealogical hypotheses—which could amount to showing that humans are pretty good at judging moral matters—but (with apologies for the repetition) this is extra work that needs to be accomplished before conservatism can supply justification in the moral realm.

Any of the epistemic justificatory theories just presented—reliabilism, coherentism, intuitionism, and conservatism—might be correct for all that I've said here. My efforts have been confined to showing very quickly how within each framework the discovery that a body of beliefs is the product of a process that the error theorist can endorse will undermine the epistemic status of those beliefs. Since plausible evolutionary hypotheses about the origins of moral thinking appear to be endorsable by the moral error theorist, moral believers must shoulder a dialectical burden: they have work to do if they are to maintain their moral beliefs in a non-dogmatic fashion.

References

Annas, J. & Barnes, J. 2000. *Sextus Empiricus: Outlines of Scepticism*. Cambridge: Cambridge University Press.

Audi, R. 2004. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press.

Ayala, F. 2010. "The difference of being human: Morality." *Proceedings of the National Academy of Sciences of the USA* 107: 9015–22.

Baillargeon, R. 2008. "Innate ideas revisited for a principle of persistence in infants' physical reasoning." *Perspectives on Psychological Science* 3: 2-13.

Bandura, A. 1999. "Moral disengagement in the perpetration of inhumanities." *Personality and Social Psychology Review* 3: 193-209.

Bartels, D., Bauman, C. Cushman, F., Pizarro, D., & McGraw, A.P. 2016. "Moral judgment and decision making." In G. Keren & G. Wu (eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making, Vol. 1*. Chichester, UK: Wiley. 479-491.

Churchland, P. 2011. *Braintrust: What Neuroscience Tells Us About Morality*. Princeton: Princeton University Press.

Darwin, C. 1879/2004. *The Descent of Man*. London: Penguin Books.

Dwyer, S. 2006 . "How good is the linguistic analogy?" In P. Carruthers, S. Laurence, & S. Stich (eds.), *The Innate Mind, Volume 2: Culture and Cognition*. Oxford: Oxford University Press. 237-55.

Enoch, D. 2011. *Taking Morality Seriously*. Oxford: Oxford University Press.

Frank, R.H. 1988. *Passions within Reason: The Strategic Role of the Emotions*. NY: Norton.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. 2008. "Strong reciprocity and the roots of human morality." *Social Justice Research* 21: 241-253.

Gould, S.J. & Lewontin, R.C. 1979. "The spandrels of San Marco and the Panglossion paradigm: A critique of the adaptationist programme." *Proceedings of the Royal Society: Biological Sciences, Series B* 205: 581-598.

Irons, W. 1996. "Morality, religion, and human nature." In W.M. Richardson & W. Wildman (eds.), *Religion and Science: History, Method, Dialogue*. NY: Routledge. 375-399.

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

Joyce, R. 2006*b*. "Is human morality innate?" In P. Carruthers, S. Laurence, & S. Stich (eds.), *The Innate Mind: Culture and Cognition*. Oxford: Oxford University Press. 257-279.

Krebs, D. 2011. *The Origins of Morality*. Oxford: Oxford University Press.

Machery, E. & Mallon, R. 2010. "The evolution of morality." In J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong, & S. Stich (eds.), *The Moral Psychology Handbook*. Oxford: Oxford University Press. 3-46.

Mates, B. 1996. *The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism*. Oxford: Oxford University Press.

Mikhail, J. 2011. *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge: Cambridge University Press.

Nichols, S. 2005. "Innateness and moral psychology." In P. Carruthers, S. Laurence, & S. Stich (eds.), *The Innate Mind: Structure and Contents*. Oxford: Oxford University Press. 353-430.

Prinz, J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Prinz, J. 2008. "Is morality innate?" In W. Sinnott-Armstrong (ed.), *Moral Psychology Volume 1: The Evolution of Morality: Adaptations and Innateness*. Cambridge, MA.: MIT Press. 367-406.

Prinz, J. 2009. "Against moral nativism." In D. Murphy & M. Bishop (eds.), *Stich and his Critics*. Malden: Blackwell. 167-89.

Rakison D. & Derringer J. 2008. "Do infants possess an evolved spider-detection mechanism? *Cognition* 107: 381-393.

Ross, W.D. 1930. *The Right and the Good*. Oxford: Oxford University Press.

Ruse, M. 1986. *Taking Darwin Seriously*. Oxford: Basil Blackwell.

Simeon, F., Regolin, L., & Bulf, H. 2008. "A predisposition for biological motion in the newborn baby." *Proceedings of the National Academy of Sciences* 105: 809-813.

Simpson, B., Harrell, A., & Willer, R. 2013. "Hidden paths from morality to cooperation: Moral judgments promote trust and trustworthiness." *Social Forces* 91: 1529-1548.

Slater, A. & Quinn, P. 2001. "Face recognition in the newborn infant." *Infant and Child Development* 10: 21-24.

Tangney, J. & Fischer, K. 1995. *Self-Conscious Emotions: The Psychology of Shame, Guilt, Embarrassment, and Pride*. New York: Guilford Press.

Tenbrunsel, A. E., & Smith-Crowe, K. 2008. "Ethical decision making: Where we've been and where we're going." *Academy of Management Annals* 2: 545-607.

Van der Linden, S. 2011. "Charitable intent: A moral or social construct? A revised theory of planned behavior model." *Current Psychology* 30: 355–374.

Wallis Budge, E.A. 1901. *Book of the Dead*. London: Kegan Paul, Trench, Trübner, and Co.

Wilson, E.O. 2012. *The Social Conquest of Earth*. London: Norton.#