8 Evolution and the Epistemological Challenge to Moral Realism

Justin Horn

Introduction

Moral realism, as I shall understand it, is the view that morality is stance-independent. That is, according to moral realism, the fundamental moral standards are true independently of the attitudes that any human beings have toward them, and independently of the attitudes that human beings would have toward them under idealized conditions of reflection (Shafer-Landau 2003, 15). As a metaethical theory of the nature of morality, moral realism has gained considerable philosophical ground over the last few decades. Recently, however, several authors have challenged moral realism on evolutionary grounds (Kitcher 2005, 2011; Joyce 2006a; Street 2006; Locke 2014). The version of the challenge that I wish to examine centers on the following claim: given the way in which our moral faculties have been shaped by evolutionary processes, those faculties cannot provide us with justified beliefs about any stance-independent moral truths (at least once we are made aware of this evolutionary influence). Insofar as the combination of moral realism and radical moral skepticism is deeply implausible, the evolutionary challenge would, if successful, give us strong reason to abandon moral realism.

Evolutionary arguments in metaethics characteristically rely on empirical assumptions, many of which are, admittedly, somewhat speculative. A complete assessment of evolutionary arguments against realism would require a detailed examination of the evidence supporting various hypotheses concerning the evolution of our capacity for moral judgment. I shall not attempt this difficult task here. Rather, in this chapter, I propose to examine the metaethical implications of one hypothesis that has been endorsed by a number of authors, both philosophers and scientists. The hypothesis is this: evolutionary forces have had a significant influence on the content of our moral judgments (see De Waal 1997, 2006; Hauser 2006; Joyce 2006a; Street 2006). Notice that this hypothesis is stronger than the plausible claim that our capacity to think morally is, in some general sense, the product of evolution. The hypothesis is committed to the stronger claim that evolution has "pushed" us in the direction of making certain moral judgments rather than others. According to the

hypothesis, just as the pressure of natural selection played a significant role in bringing it about that tigers have sharp claws (rather than dull ones) and that zebras are speedy (rather than slow), so too did the pressure of natural selection play a significant role in bringing it about that human beings have a very strong tendency to regard certain things as morally valuable or morally reprehensible. Examples of evolutionarily favored moral attitudes might include the widespread positive moral regard enjoyed by activities such as caring for one's own children and reciprocating benefits provided by others, and the negative moral regard commonly held for defecting from agreements or casually harming one's kin. There are, of course, variations in the precise form that such "moral regard" takes in different individuals and cultures. But it is hard to deny that even these vague generalizations get at real patterns in moral judgment, patterns for which there must be some causal explanation. The hypothesis in question holds that some such deep patterns in moral judgment are significantly due to the pressure of natural selection, such that had we evolved differently, we would make moral judgments with very different content.

I hasten to note, however, that the hypothesis is not intended to serve as a complete explanation of why we make all of the particular moral judgments we do. The truth of the hypothesis is consistent with the fact that there are many other significant influences on the content of our moral judgments – influences that include rational reflection, as well as a variety of social, cultural, and historical factors. The claim is only that evolution has been one powerful causal influence on the content of our moral judgments.

To avoid confusion, I should note a few of the consequences of the hypothesis. First, because the hypothesis is not intended to explain all moral phenomena, it is not threatened by the existence of certain moral phenomena for which the best explanation is not an evolutionary one – for example, the changes in moral attitudes toward women and racial minorities that have taken place in the United States over the last 150 years. The following analogy might be helpful here. We have excellent reason to believe that genetic factors have a significant influence on our personalities. This belief is in no way threatened by the existence of personality-related phenomena for which the best explanation invokes nongenetic factors, such as a case in which identical twins have substantially different personalities. Such cases merely show that genetic factors are not the only causal factor contributing to personality development. Similarly, certain moral phenomena may have a non-evolutionary explanation, even if the influence of evolution on our moral beliefs is pervasive.

Furthermore, the claim that evolution has significantly influenced the content of our moral judgments does not entail that we will inevitably make all of the moral judgments that would be adaptive. Nor does it entail that all of

the judgments we do make will be adaptive. On the most likely model of the development of our moral capacities, evolution has endowed us with certain very general evaluative dispositions regarding harm, fairness, purity, and the like (see de Waal 1997, 2006; Kitcher 2005, 2011; Street 2006; Haidt and Joseph 2007; Haidt 2012). These dispositions in turn shape the content of our moral judgments significantly, such that if we had a different set of general evaluative dispositions, we would come to very different conclusions about a variety of moral matters. Nonetheless, these general evaluative dispositions do not wholly determine which moral judgments we make; rather, they can be channeled in a variety of ways by culture, learning, and rational reflection. Thus, the foregoing model is consistent with us making moral judgments that are in fact quite harmful to our fitness, such as the judgment that I am required to sacrifice my life for my country.

With these clarifications in place, I will henceforth assume for the sake of argument that the hypothesis is true. The challenge for the realist is to provide some defensible account of the relationship between the putative stance-independent moral truths and the evolutionary pressures that (we are assuming) have shaped our moral judgments. I'll describe three general accounts of this relationship and argue that the realist's best hope lies in establishing what I'll call the indirect tracking account. I'll then argue that any attempt to establish such an account will rely on premises for which the recognition of evolutionary influence will have already defeated our epistemic justification. In light of this defect, I argue that if our initial hypothesis is true, moral realists are saddled with the conclusion that all of our positive moral beliefs – those beliefs that attribute moral properties – are unjustified. Since this sort of radical skepticism about morality is implausible, I conclude that if evolutionary forces have strongly influenced the content of our moral judgments, then we have strong reason to reject realism as a metaethical position.

The Distortion Hypothesis

If we assume that evolutionary forces have "pushed" us in the direction of making some moral judgments rather than others, the realist must give some account of the relationship between this evolutionary influence and the putative stance-independent moral truths. One possibility is that there is no positive correlation between the stance-independent moral truths and those moral judgments that evolution has "pushed" us toward. Let us call this the distortion hypothesis. On the face of it, the distortion hypothesis seems to represent a worst case scenario for the realist. If the distortion hypothesis is true, then evolutionary pressures are no more likely to have pushed us toward true moral beliefs than, say, were we to have based our moral beliefs on a random drawing from a hat containing all logically possible moral judgments.

The hat-drawing method is obviously very unlikely to consistently yield true moral beliefs. But if the distortion hypothesis is true, one considerable influence on the content of our moral judgments is, in all relevant respects, exactly like hat-drawing. If evolutionary influences on our moral judgments are sufficiently deep, and we have no way of correcting this distorting influence, then many of our moral beliefs are very likely to be false.

According to the hypothesis under consideration, evolutionary influences on our moral beliefs are in fact quite deep. Although certainly other factors may influence our moral judgments, the hypothesis suggests that many of our central moral commitments are largely the result of evolutionary pressures. Thus if this hypothesis is true, a realist seeking to embrace the distortion hypothesis without succumbing to skepticism must therefore argue that we possess the tools for weeding out and correcting even deep and widespread errors among such moral judgments. One way to argue for this conclusion would be to postulate a special faculty of moral intuition. It is not clear that this would solve the problem, however. Unless the realist also holds that we can reliably distinguish the outputs of our special faculty of intuition from those moral judgments that have been conditioned by evolutionary forces, the most that a faculty of intuition will achieve is to mix some true beliefs in with those mostly false beliefs that we have due to evolutionary pressures. Without any tool to reliably separate the two, a large percentage of our moral beliefs will still be false.

Some have argued that performing such a separation is relatively easy – we can simply look at the content of a particular judgment to figure out whether it is likely to have been influenced by evolutionary pressure. This is so, it is argued, because "biological evolution would be expected to produce a bias toward favorable evaluations of things that promote one's own inclusive fitness; intuitions that do not imply favorable evaluations of things that promote one's own inclusive fitness are not candidates for being products of this particular bias" (Huemer 2008, 381; see Shafer-Landau 2012, 5–8). This method of sorting is problematic, however. First, since evolutionary change takes time and environments can change rapidly, natural selection can produce organisms with traits that no longer promote their own inclusive fitness. Consider, for example, human tastes in food. As it happens, most of us find ourselves attracted to sweet, salty, and fatty foods. In environments where food is plentiful, these preferences can be quite maladaptive, leading people to consume more calories than would be healthy, and often leading to early death. From the point of view of survival, it would probably be far better for those of us living today if we were to crave whole grains and leafy green vegetables, and to find large quantities of red meat repulsive. Nonetheless, there is a straightforward and plausible evolutionary explanation of our cravings for fats and sugars: in the ancestral environment, in which food was scarce and starving

was a very real threat, it was adaptive to be motivated to consume the most calorific foods available. We inherited from our ancestors a palate that helped them survive but that often leads us today to heart problems and obesity. Thus natural selection can produce traits that are, in our current environment, quite detrimental to our own inclusive fitness.

Furthermore, it is not plausible that evolution influenced the content of our moral judgments by directly selecting for some particular judgments over others. On the contrary, the influence of evolution was likely much more indirect: certain general evaluative tendencies were selected for over time, and these evaluative tendencies in turn strongly influence which moral judgments we end up making. Given this sort of influence, it wouldn't be surprising if such dispositions sometimes "misfired" to produce particular judgments that turn out not to be reproductively advantageous, especially given significant differences between the ancestral environment and our own. Consider just one (admittedly speculative) example. It clearly promotes one's own inclusive fitness to ensure that one's own offspring – and to a lesser extent the offspring of one's close relatives – survives to reproductive age. If, in the ancestral environment, most of the small children that one came across were closely related to oneself, a standing disposition to be gentle toward all small children and refrain from harming any of them would be highly adaptive, and could in principle be selected for. In our present environment, we often come across small children to whom we are not related. Perhaps, in some circumstances, harming them would promote one's inclusive fitness. Nonetheless, there could very well be an evolutionary explanation for why we are strongly disposed to regard harming small children as forbidden, even when it would be advantageous to harm them.

To be clear, I am not attempting to offer a complete and accurate explanation of our attitudes toward small children. Rather, the example demonstrates that the mere fact that a judgment does not currently promote one's own inclusive fitness is consistent with the hypothesis that the judgment is substantially the product of natural selection. Furthermore, to the extent that evolutionary influence on our moral judgments took place largely at the level of very deep and general evaluative dispositions (for example, by inclining us to regard certain very general features such as harm, fairness, loyalty, and purity as having positive moral relevance), it will be very difficult to find substantive moral judgments that are plausibly entirely isolated from such dispositions, even when we consider judgments that happen to be detrimental to our fitness in the present environment.

Even if we cannot identify exactly which moral beliefs have been shaped by evolutionary forces, one might think that the widely endorsed method of reflective equilibrium would allow us to correct for any potentially distorting effects that evolution has had on our moral judgments (Rawls 1971). Perhaps

by "testing" our judgments about moral principles against our judgments about particular cases and vice versa, while also seeking coherence between our moral judgments and background theoretical considerations, we can root out even deep moral errors generated by evolutionary pressure.

If the distortion hypothesis is true, however, it is doubtful that such a method will be of much help. The method of reflective equilibrium essentially involves simply working back and forth between judgments about moral principles and judgments about particular cases, adjusting each in light of the other until an adequate degree of coherence is achieved. If the set of initial moral judgments with which we begin inquiry is sufficiently corrupt, however, such a process of mutual adjustment is unlikely to be promising as a way of arriving at a stance-independent moral truth. As Sharon Street points out, if the distortion hypothesis is correct, then this sort of reasoning will simply involve "assessing evaluative judgments that are mostly off the mark in terms of others that are mostly off the mark" (Street 2006, 124). Thus, it seems that if the distortion hypothesis is true, a large percentage of our moral beliefs are very likely to be false, even if we prune them so as to bring them into reflective equilibrium.

The distortion hypothesis is therefore an unattractive option for any realist who believes that some of us are reliable moral judges. For this reason, realists might be encouraged to notice the following peculiar feature of that hypothesis: it would be very difficult to establish that the hypothesis is actually correct. To establish that there is indeed no positive correlation between the moral truth and the evaluative judgments that were selected for, it seems that we would need to compile a rough list of some moral truths and then compare the moral truths to those evaluative judgments favored by natural selection. (Elliott Sober makes a similar point in Sober 1994, 107.) Only by having some information about the contents of each list could we provide evidence that these contents were not correlated. If we possessed the information required for this task, however, then clearly any skeptical argument would be very hard to get off the ground. After all, in such a situation, we would already have a rough list of at least some moral truths. Much like global skepticism, it seems that the distortion hypothesis is not one that can be coherently asserted with confidence.

Yet it remains a troubling possibility, for an obvious reason. If the distortion hypothesis is correct, very many of one's moral beliefs are likely to be false, since there is no correlation between those moral beliefs that natural selection has pushed us toward (and thus, many of the moral beliefs that humans tend to have) and the moral truth. It seems to be a very plausible epistemological principle that if one has undefeated reason to think that one's beliefs in a domain have a high probability of being false, one cannot be justified in holding those beliefs. Thus, if realists have reason to believe that there is even a fairly high probability that the distortion hypothesis is correct, realism faces a

serious epistemological challenge. Of course, I have not yet given any reason to believe that there actually is a high probability that the distortion hypothesis is correct. At this point, the thing to notice is merely that the distortion hypothesis could, in principle, threaten to undermine the justification of our moral beliefs even if it cannot be firmly established as correct.

The Direct Tracking Hypothesis

Realists might want to forestall this skeptical possibility by arguing that the evolutionary influence on our moral beliefs has been largely benign. One way of doing so would be to suggest that evolutionary pressures have pushed us toward the stance-independent moral truth, precisely because it was adaptive for our ancestors to grasp the moral truths in question. Let us call this the direct tracking hypothesis. Such a hypothesis, if true, would not only save the realist from epistemological objections based on evolutionary grounds but also would in fact provide the realist with a powerful tool for defending our general moral reliability.

This hypothesis is unacceptable on scientific grounds, however. In particular, it is inferior to a competing hypothesis, which Sharon Street calls the adaptive link account. According to the adaptive link account, "tendencies to make certain kinds of evaluative judgments rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent evaluative truths, but rather because they forged adaptive links between our ancestor's circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous" (Street 2006, 127).

The main problem with the direct tracking account is that the most promising explanations of the evolutionary influence on the content of our moral beliefs simply needn't make any reference to the existence of moral facts. Indeed, it's not clear how postulating such facts would contribute anything to such an explanation. In contrast, consider the best explanation of the origins of our capacity for detecting mid-sized physical objects. Any acceptable explanation of our perceptual abilities will invoke the fact that non-veridical perceptions of mid-sized physical objects (say, predator or prey) would tend to be detrimental to the fitness of an organism. If an organism tends to form beliefs to the effect that it is being chased by predators when this is not so, it will end up wasting a lot of valuable time and energy running and hiding. Still worse, if an organism tends not to form beliefs that it is being pursued by a predator on those occasions when it is in fact being pursued, that organism's genes are likely to be swiftly removed from the gene pool. In short, when it comes to avoiding predators, the truth of one's perceptual beliefs is of paramount importance.

In contrast, it is not at all clear how the truth of one's moral judgments can play any analogous role in an evolutionary explanation of our moral abilities. Other things being equal, it seems it would be adaptive for an organism to believe that it ought to take care of its offspring, and maladaptive to believe that it ought to kill them. But the adaptiveness (or lack thereof) of these judgments would remain exactly the same if it were to turn out, quite surprisingly, that we have a fundamental moral obligation to kill our own offspring. In morality, it simply does not seem that beliefs are ever adaptive in virtue of being true. Thus, we should expect selection for moral judgments that form adaptive links between circumstances and behavior, regardless of whether such beliefs are true or false.

One should note the limitations of the preceding remarks. I have not argued that the mere fact that stance-independent moral facts play no role in scientific explanations justifies eliminating them from our ontology. The present claim is much more modest. Given that we can explain everything worth explaining about the evolutionary influences on moral judgment without postulating moral facts, considerations of parsimony give us a reason to prefer the adaptive link account to the direct tracking account. Thus, we may conclude that while the direct tracking account would save the realist from a skeptical conclusion, it is unacceptable on scientific grounds.

Indirect Tracking and Preestablished Harmony Explanations

I've argued that both the distortion hypothesis and the direct tracking hypothesis are unpromising ways for the realist to reconcile the notion that we have justified beliefs about stance-independent moral truths with the (putative) fact of evolutionary influence on the content of our moral judgments. There is another possibility, however. It may be that there is a strong correlation between the stance-independent moral truths and those moral judgments that were selected for, such that the moral judgments that were selected for are mostly true, but were not selected for because they were true. I say "mostly" because the realist needn't insist that evolutionary pressures have pushed us toward the truth in every case to surmount the epistemological challenge. As David Copp points out, the realist can resist a skeptical conclusion, provided that "our beliefs tend to do well enough in tracking the moral truth that rational refection can in principle correct sufficiently for any distorting influence" (Copp 2008, 194). The position we must consider, then, is one that accepts the adaptive link account as an explanation of why certain moral judgments were selected, while still holding that the moral judgments selected are close enough to the truth. I think that this view, which I will call the indirect tracking hypothesis, is the most promising avenue for the realist.

How might one defend the hypothesis that true moral beliefs were not selected for because they were true, but that nonetheless evolutionary influences have pushed us toward mostly true moral beliefs? The most promising explanation appeals to the widely accepted principle that any moral facts that exist supervene on natural facts: natural facts fix the moral facts in the sense that, necessarily, any two states of affairs that are exactly alike in all natural respects must be exactly alike in all moral respects. According to the evolutionary hypothesis presently under consideration, evolutionary forces have pushed us toward the acceptance of moral beliefs that are appropriately related to certain natural facts (namely, facts about survival and reproduction). If the natural facts that our moral beliefs tend to track are systematically related to the moral facts, this opens the door for a "pre-established harmony" explanation of the correlation between the moral judgments that were selected for and the realist's stance-independent moral truths (Enoch 2010; Skarsaune 2011).

Suppose, then, that the realist accepts that certain moral beliefs were selected for, as described by the adaptive link account. The realist might proceed to argue that these moral beliefs are (mostly) true, because the features that moral judgments were selected to track either constitute or closely correlate with moral features. We can explore how such a strategy would work by considering a simple form of naturalistic realism: hedonistic utilitarianism.

The hedonistic utilitarian might admit that moral beliefs were selected not for their truth, but for their tendency to motivate individuals to behave in ways that increased reproductive success. But the utilitarian might then claim that the moral beliefs that evolution has conferred on us are for the most part reliable. The utilitarian needn't simply see this as a convenient coincidence, but could argue for it as follows. Pleasure is intrinsically good and pain is intrinsically bad. Given this, one can imagine why natural selection would, to a considerable degree, favor true moral beliefs rather than false ones. After all, pain is typically an indicator of bodily harm, so organisms that tend to view pain as bad would tend to survive longer than those who do not. Likewise, pleasure is often an indicator of bodily benefit (or in the case of sexual pleasure, of reproductive success), and therefore organisms that see pleasure as good would tend to have greater reproductive success than those that do not. Thus, while evolutionary forces may have led us astray in some cases (for instance, the widespread belief that we have only very weak obligations to distant strangers), it is no accident that it has given us mostly true moral beliefs.

I use utilitarianism as an example, but it is important to note that this sort of explanatory strategy could in principle be used for a wide variety of normative theories. It needn't be limited to reductive accounts, or even to

naturalist accounts. Any view that claims that the moral facts supervene on natural facts could in principle tell this sort of story, by first linking certain natural features of the world with moral features, and then arguing that it was (for the most part) adaptive for our ancestors to regard those natural features as good, even though the explanation of why this is adaptive makes no reference to the truth of their judgments.

The nonnaturalist realist David Enoch adopts this sort of strategy to respond to Street's Darwinian Dilemma. According to Enoch, what I have called the indirect tracking hypothesis can be adequately supported if we merely accept that "survival or reproductive success (or whatever else evolution "aims" at) is at least somewhat good" (Enoch 2010, 430). This claim is not intended as a reductive account of what goodness is; it is merely a rough and ready claim that in most circumstances, survival has value. Enoch argues that if survival has value, and viewing survival as valuable was selected for, then evolution might have left us with mostly true moral beliefs, even if the truth of these moral beliefs plays no role in the explanation of why they were selected for.

It is not clear that such a modest assumption is sufficient to explain the correlation Enoch aims to explain. The claim that survival is at least somewhat good is compatible with the claim, for instance, that the beauty of nature is of far greater value and that we are all obligated to sacrifice our own survival to maximize natural beauty. Likewise, Enoch's normative claim is compatible with the view that while survival is good, this goodness is outweighed by the goodness of excruciating suffering. An indefinite number of logically possible, internally coherent ethical systems are compatible with the claim that survival is at least somewhat good, and many of these systems differ dramatically from our moral intuitions in far-reaching, systematic ways. Thus, even assuming that survival is somewhat good, the realist still needs an explanation of why our system of intuitive moral judgments (which incorporates this assumption) approximates the stance-independent moral truth, while all other such internally coherent sets incorporating it do not.

In general, though, indirect tracking accounts seem attractive because they have the potential to provide an explanation of a correlation between those moral beliefs favored by natural selection and the stance-independent moral truth, and all this without giving up the scientifically preferable adaptive link hypothesis. Nonetheless, such accounts suffer from a serious defect. When presented with a claim linking the moral to the nonmoral, we are entitled to ask what evidence or justification is on offer for the claim. The realist answer, it seems, will typically rely on substantive normative ethical views. This was clearly the case in the previous utilitarian example, as well as the case of David Enoch's more modest bridge principle. In a similar vein, Erik Wielenberg attempts to vindicate our moral judgments in the face of

evolutionary influence by assuming the normative claim that there are "moral barriers" that surround all creatures with sufficient cognitive capacities. These cases do not seem to be exceptions to the rule. As realist David Brink writes, ''determination of just which natural facts and properties constitute which moral facts and properties is a matter of substantive moral theory" (Brink 1989, 177–178). The problem is that invoking substantive normative ethical views at this point in the dialectic begs the question, since the reliability of these views is exactly what is at stake.

The question for the realist is whether evolutionary influences have left us with the capacity to form moral beliefs that (at least roughly) track a stance-independent moral truth. Supposing we have ruled out the direct truth tracking account, we are left with two options: the indirect tracking hypothesis or the distortion hypothesis. If the distortion hypothesis is correct, then most of our intuitive moral judgments are false. If the indirect tracking hypothesis is correct, then a large number of our intuitive moral judgments are true, at least enough such that rational reflection could (in principle) weed out the bad apples. The trouble for the realist is this: how do we figure out which of these two possibilities obtains?

If we are at all unsure, it simply will not do to invoke substantive normative judgments at this point. Consider the following analogy. Suppose you discover that you've been brainwashed by a cult leader, who has given you all sorts of supernatural beliefs, which are based on visions he experienced while taking a brand new Miracle Drug. Further suppose that you are genuinely unsure whether Miracle Drug visions are a reliable guide to the supernatural truth, and you are trying to ascertain whether or not this is so. Clearly it would not do to "test" the beliefs that the leader formed when using Miracle Drug against your own convictions about the supernatural. After all, you know that your beliefs about the supernatural are the result of the cult leader's brainwashing, so of course his supernatural beliefs will pass this "test," whether Miracle Drug visions are reliable or not. (For a similar point, see Copp 2008, 197.)

Analogous things could be said about the evolutionary influences on our moral beliefs. If we are trying to determine whether evolutionary forces have pushed us toward the moral truth (as the indirect tracking hypothesis says) or not (as the distortion hypothesis says), it will be of no use to "test" the moral beliefs that would be selected for against our intuitive moral judgments. For we know (or so we are supposing) that our moral judgments have been heavily shaped by evolutionary forces. For this reason, the moral beliefs that have been selected for would be very likely to pass this test, even if the distortion hypothesis were correct.

Is there some way of establishing the indirect tracking hypothesis that does not rely on any first-order normative views? This would be nice for the realist,

but it does not seem promising. If we accept some version of Hume's dictum that moral claims cannot be established by arguments that invoke no moral premise whatsoever, we must admit that any attempt to establish the indirect tracking hypothesis will rely on normative judgments. I've argued that relying on normative judgments to establish the truth of the indirect tracking hypothesis over the distortion hypothesis begs the question. Thus, any argument for favoring the indirect tracking hypothesis over the distortion hypothesis will be question begging.

Moral Realism and Skepticism

I've argued thus far that arguments for the indirect tracking hypothesis are question begging, while the distortion hypothesis cannot be coherently asserted confidently. One might be tempted to conclude that without any way of resolving which hypothesis is the correct one, the epistemological challenge to moral realism flounders. After all, given everything that I've said, perhaps the indirect tracking hypothesis is correct and our epistemological situation is pretty good. So nothing I've said can be thought to undermine realism.

Furthermore, one might press the following line of argument. Suppose one were to call into question the justification of our perceptual judgments by challenging us to show that they themselves were not distorted in some deep way. One natural reply to such a challenge is to point out that the most plausible account of our basic perceptual capacities will be (to a significant extent) a direct tracking account, according to which the ability of those capacities to yield true judgments was essential to their being selected for. And this reply does seem adequate to vindicate our perceptual capacities to some degree. But notice: we can only establish a direct tracking story about the evolutionary origins of our perceptual capacities by relying on those capacities from the outset. Without the input of sensory observations, scientific theorizing about the nature of evolutionary influence on our perceptual capacities could never get off the ground (see Schafer 2010 and Vavova 2014). And yet we do think that we are justified in believing things on the basis of our senses. So it seems plausible to suppose that our perceptual judgments have some justification from the very beginning.

The moral realist might insist that similar considerations allow us to justifiably accept the indirect tracking hypothesis over the distortion hypothesis in the case at hand. If we are willing to grant some justification to our perceptual judgments from the outset, there seems to be no reason not to allow that our moral judgments enjoy a similar degree of justification at the outset of inquiry. Once we grant this, though, it seems that the realist can rely on her

justified moral beliefs to rule out the distortion hypothesis and find in favor of an indirect tracking view (see Schafer 2010).

This is an elegant line of argument, but I think it can be resisted. We should grant the first point: we should, at the outset of inquiry, regard our intuitive moral judgments as having some modest degree of justification. The question we must ask is whether this justification is undercut by the time we face the question of whether to prefer the indirect tracking hypothesis to the distortion hypothesis. And it seems to me the answer is yes.

The first thing to notice is that in seeking an evolutionary vindication of our perceptual judgments, there is never a moment at which we have an explanation of the origins of our perceptual faculties that completely leaves open the question of whether they are reliable in tracking stance-independent facts about our surroundings. We begin with perceptual judgments that have some degree of justification, we do a lot of scientific inquiry, and we wind up with additional reasons to trust our perceptual faculties: our best explanation of their emergence vindicates their (approximate) reliability. But the second thing to notice is that we can imagine things being different. And in such imagined cases, the wrong kind of genealogy of our perceptual judgments could undermine their justification, even while leaving it open whether or not such judgments were actually reliable.

Imagine you were to discover something shocking about your perceptual judgments: they are never caused by external physical objects. Further, imagine you discover this in a manner completely independent of your perceptual capacities – perhaps God directly imparts this knowledge to you. It turns out that all of your perceptions are constantly being caused directly by some supernatural creature. This supernatural creature is akin to Descartes's evil demon, with one important difference: we have no idea whether he is evil. (For some reason, God neglects to tell us this part.) Indeed, we have no indication of the being's intentions whatsoever. Call this creature the Demon of Unknown Intentions (DUI).

With this new and disturbing information about your perceptual capacities, you start to worry about your ability to reliably form beliefs about external physical objects. You reason as follows: on the one hand, it's consistent with your newfound knowledge that the DUI is benevolent and only gives you veridical perceptual experiences. Perhaps you only have the experience of a tree when there is indeed an external physical tree in your vicinity. Perhaps, in fact, the demon is necessarily benevolent, and so couldn't possibly deceive you in any deep and undetectable way. On the other hand, it's also entirely consistent with your newfound knowledge that the DUI is entirely deceiving you. Perhaps, as far as physical reality goes, you are just a brain in a vat, or an eight-armed slimy creature, or perhaps there is no external physical world at all.

What would be reasonable to conclude if (somehow) you were to learn that, as a matter of fact, all of your perceptual experiences were caused by the DUI? You could hope for a kind of indirect tracking explanation. Perhaps you could find some central regularities in the world of your experience and assume that these correspond to physical reality (perhaps relying on our a priori entitlement to trust our perceptual capacities to justify this claim) and then deduce that your initial perceptual beliefs were close enough to veridical to correct any distorting influence through reasoning. For example, you might note that the DUI has made your experiences such that Newtonian mechanics seems roughly true of macroscopic objects. Since, you insist, Newtonian mechanics is roughly true of such objects, the DUI has probably not led you too far astray.

But here this reply seems totally unconvincing. Once you learn that your perceptual judgments are caused by something wholly distinct from any physical objects they seem to report, something that you have no independent reason to regard as a reliable source, the initial justification provided by your perceptual judgments is defeated. Absent any other way of finding out about a world of objectively existing external physical objects, it seems that all of your beliefs about them would be rendered unjustified.

This remains so even if we weaken the case a bit, so that the DUI is not wholly responsible for your perceptual beliefs. Suppose you are informed that the DUI is only one significant influence on the content of your perceptual judgments. Nonetheless, you learn, (a) this influence is such that you have no way of isolating any perceptual judgments that are known to be free of the influence of the DUI, and (b) the influence of the DUI is sufficiently powerful that the following is true: had the DUI influenced you differently, you would make radically different perceptual judgments. It seems to me that learning of even this more modest influence of the DUI on your perceptual judgments has deep skeptical consequences. Continuing to believe that your perceptual beliefs accurately represent an external physical reality in such a case requires trusting that the influence of the DUI has pushed you toward, rather than away from, the truth. But this is exactly what you have seem to have no reason to believe, and no way of figuring out.

In the case of the DUI, what defeats our initial justification for our perceptual beliefs is that we justifiably accept an account of the origins of those beliefs that (a) rules out a direct tracking explanation and (b) gives us no reason to prefer an indirect tracking account to a distortion account. Once we have this, the initial warranted confidence we had in regarding those beliefs as faithful representations of an external physical reality disappears. Yet, if our initial hypothesis about evolutionary influence on the content of our moral judgments is true, then – at least once we come to realize its truth – we seem to be in an analogous epistemic situation. For we will have identified

one deep influence on the content of our moral judgments, where the best explanation of the nature of this influence (a) rules out a direct tracking account and (b) gives us no reason to prefer an indirect tracking account to a distortion account. Once we have this in hand, it seems that – at least insofar as we regard our moral beliefs as attempts to represent a stance-independent moral reality – all of our moral beliefs will be unjustified.