Arguments from moral disagreement to moral skepticism
Richard Joyce

Penultimate draft of paper appearing in D. Machuca (ed.),

1. Introduction

References to the phenomenon of moral disagreement appear conspicuously in several areas in metaethics. One well-known argument—J. L. Mackie’s (1977) argument from relativity—seeks to establish on the basis of widespread disagreement that moral judgments are false. Another argument—whose general form goes right back to classical philosophers like Agrippa and Sextus Empiricus—seeks to establish on the basis of widespread disagreement that moral judgments lack justification. Both conclusions can properly be called versions of moral skepticism. If moral skepticism is the view that there is no moral knowledge, then both the former error-theoretic position that moral judgments express false beliefs, and the latter epistemological position that moral judgments express unjustified beliefs, are forms of skepticism. (The noncognitivist position that moral judgments do not express beliefs at all is also a form of moral skepticism.)

This chapter is almost exclusively diagnostic in ambition. I aim to sketch out the complex structure and interrelations of these skeptical arguments based on moral disagreement, without advocating any of them. Though I am very sympathetic to moral skepticism, I am yet to be convinced that there is an argument for it based on the phenomenon of moral disagreement that I would find persuasive in the absence of that sympathy. There are a couple of recurring themes that are worth highlighting in advance. One is the fact that the debate frequently hinges on empirical matters; the other is the complicated relation between skepticism and moral naturalism: often, as we shall see, the skeptic is opposed to naturalism, but other times the possibility of a defensible moral naturalism turns out to be a skeptic-friendly result.

2. Mackie’s argument from disagreement

Mackie’s argument from relativity is poorly named, since that title might be taken to suggest that he is seeking to establish moral relativism, which he is definitely not. I prefer to call it “the argument from disagreement.”¹ As a first stab at the argument, consider this:

¹ Others have this preference also. See Brink 1984; Loeb 1998. Charitably, one might assume that Mackie has in mind so-called descriptive relativism, but since I think that that is also a misleadingly labeled thesis, the charity is limited!
ARGUMENT FROM QUANTITY OF DISAGREEMENT:
P1: If moral realism were true, then we would observe no more than such-and-such amount of moral disagreement in the world.
P2: In fact, we observe a great deal more moral disagreement than such-and-such.
C: Therefore, moral realism is false.

Even if we could correct for the glaring imprecision of “such-and-such,” both premises are vulnerable to realist objections. Realists object to P2 by trying to downplay the amount of moral disagreement we actually observe. Much of what we might think of as moral disagreement, they say, is really disagreement over non-moral beliefs masking more fundamental moral agreement. And realists cast doubt on P1 by pointing out that it is not clear how much convergence or divergence in moral opinion moral realism really predicts. They seek partners in innocence: other domains where there is a great deal of disagreement (e.g., competing scientific hypotheses) but for which we are very disinclined to reach for an anti-realist conclusion.

In response, anti-realists can point out that it’s not just the amount of disagreement that’s the issue; it’s the nature of it. Moral disagreement, they might say, is characterized by an unusual kind of intractability, persistence, emotiveness, and insensitivity to evidence. In fact (they might add), it’s often not even clear what would count as evidence for the truth of one moral judgment as opposed to a contrary one (Harman 1977). And perhaps the disagreement that attends the partners in innocence, widespread though it may be, lacks these qualities. So the argument shifts focus:

ARGUMENT FROM QUALITY OF DISAGREEMENT:
P1: If moral realism were true, then moral disagreement would not be so intractable, persistent, emotive, and insensitive to evidence (etc.).
P2: Moral disagreement is so intractable, persistent, emotive, and insensitive to evidence (etc.).
C: Therefore, moral realism is false.

Again, realists may object to P2 by claiming that there is more hope of convergence in moral disputes than Mackie suggests. If many moral disagreements are really at bottom disagreements over non-moral matters, then perhaps they are not so intractable after all. Realists can cast doubt

---

2 One way of plausibly denying P1 is to draw attention to the fact that realism is compatible with forms of relativism, and relativism does not predict convergence in moral opinion. On the other hand, moral relativism has some trouble accommodating the existence of disagreement. (If when I say “X is wrong” I mean from point of view φ, and when you say “It is not the case that X is wrong” you mean from point of view ψ, then what appears to be a disagreement turns out not to be one.) So the possibility of relativistic realism complicates matters considerably, and for this reason I will bracket it off from this chapter. Besides, it doesn’t appear to be a possibility that is on Mackie’s radar; he seems to think of moral realism as necessarily an absolutist position. (Perhaps, like many, he confuses objectivism and absolutism.) This seems to me a flawed taxonomy (see Joyce 2015), but in this chapter I’ll accept it for the sake of argument.
on P1 by pointing out that it is not clear how much intractability, etc. of moral disagreement would be predicted by moral realism; perhaps moral realism is compatible with moral disagreement’s having these qualities. At this point, Mackie deploys an argument with the form of inference to the best explanation. The phenomenon to be explained is moral disagreement (“the well-known variation in moral codes from one society to another and from one period to another, and also the differences in moral beliefs between different groups and classes with a complex community” [1977: 36]), bearing in mind both its quantity and aforementioned qualities. Mackie compares two explanatory hypotheses. First there is the realist hypothesis: that moral codes “express perceptions, most of them seriously inadequate and badly distorted, of objective values” (37). As for the second hypothesis—the one which Mackie prefers—we shall have to have some discussion of its content. At the very least, it involves the claim that:

i) moral codes “reflect people’s adherence to and participation in different ways of life” (36).

On this reading, though, it is difficult to see how an anti-realist conclusion is supposed to follow, even if we have grounds for thinking that the second hypothesis is correct, for the truth of (i) doesn’t obviously exclude the truth of the realist hypothesis.3 Indeed, (i) looks rather like a friendly supplement to the realist hypothesis: the realist hypothesis stating that many of our moral perceptions are “badly distorted,” and (i) then explaining why this is so. In light of this, it is tempting to read the second hypothesis as requiring an additional claim:

ii) … and there are no objective moral facts.

On the realist’s hypothesis, then, there are objective moral facts but because we disagree so much we must not be very good at accessing those facts; and on the anti-realist’s hypothesis there are no objective moral facts and our disagreement is the result of our essentially “making up” morality to suit various practical needs that differ among cultures and individuals. And then Mackie invites us to agree with him that the second explanatory hypothesis is much more plausible.

One thing that should be noticed about this argument is that even if Mackie were to succeed in establishing the second hypothesis over the first, he still falls short of his ultimate metaethical conclusion: error theory. Error theory is not the view that there are no objective moral facts; it is the view that there are no moral facts simpliciter. Even if one agrees that there are no objective moral facts, one might well embrace a non-error-theoretic view that there are non-objective moral facts (e.g., some form of constructivism). To exclude this possibility, Mackie might countenance an even stronger version of the anti-realist hypothesis, one that combines (i) with:

iii) … and there are no moral facts.

3 As Folke Tersman has pointed out (2015a, 2015b). See also Enoch 2009: 22.
The reason, I think, that Mackie is not too bothered in this context about the difference between (ii) and (iii) is that by the time that he puts forward the argument from disagreement he has already, to his own satisfaction, established that objectivity is an essential feature, conceptually speaking, of moral properties. Thus the difference between (ii) and (iii) is rather like the difference between “There are no four-sided squares in the box” and “There are no squares in the box.” Since squares are necessarily four-sided, showing that there are no four-sided squares suffices to show that there are no squares simpliciter; likewise, if moral facts are necessarily objective, then showing that there are no objective moral facts suffices to show that there are no moral facts simpliciter.

One might point out that (i)+(iii) still doesn’t entail the moral error theory, since it is compatible with noncognitivism. The error theorist and the noncognitivist agree that there are no moral facts; where they disagree is that the former maintains that moral speakers attempt to state moral facts, whereas the latter holds that moral discourse was never in the fact-stating business to begin with. But, again, by the time Mackie gets to presenting the argument from disagreement in his 1977 book, he has already put forward arguments against noncognitivism. Thus, by establishing the (i)+(ii) hypothesis over the realist hypothesis, Mackie thinks he is establishing the moral error theory, but only with the help of arguments that have come earlier in his chapter.

The problem with adding (ii) or (iii) to the second hypothesis, however, is that doing so seriously undermines the grounds we might have for endorsing it. The claim made by (i) is at bottom an empirical claim. Cross-cultural investigation might reveal that societies do indeed construct their moral codes to suit their circumstances. It might reveal, for example, that societies that permit polyandrous marriage arrangements do so because of some unusual environmental feature, such as a paucity of farmable land (see Starkweather & Hames 2012). It might reveal that large-scale and complex societies are more likely than smaller and simpler societies to endorse and enforce fairness norms governing interactions with strangers (see Ensminger & Henrich 2014). And so on. In principle, then, we might muster evidence in support of (i), but none of this evidence would support (ii) or (iii). This raises the question of why anyone, in attempting to explain moral disagreement, would prefer to maintain either of the hypotheses (i)+(ii) or (i)+(iii) over the less committed hypothesis (i). After all, the hypothesis that denies objective moral facts or denies moral facts will be worse off than the agnostic hypothesis if it turns out that there are phenomena other than moral disagreement whose explanations do require the existence of these facts (a possibility that, if we are seeking only to explain moral disagreement, we have no grounds for excluding). Moreover, the additional clauses (ii) or (iii) won’t serve to better explain any aspect of the phenomenon of disagreement. In fact, what business does an explanation have in denying the existence of something? Compare the perfectly reasonable claim “The best explanation of moral disagreement remains silent on the existence of wombats” to the bizarre claim “The best explanation of moral disagreement denies the existence of wombats.” (Note that the contrast remains if we replace mention of wombats with mention of something in which we don’t believe—unicorns, say.)
It becomes clear, then, that the active denials embodied in (ii) and (iii) should not be considered elements of the explanatory hypothesis. The explanatory hypothesis is simply proposition (i)—though of course (i) as it is worded here is merely a stub that stands in for a much more complex account—which is silent on the existence of both objective moral facts and moral facts *simpliciter*. But the problem remains that even if (i) is the best explanation of the moral disagreement, it is entirely compatible with moral realism. We might at this point declare the argument from disagreement a flop. Or we might try to salvage a case for the error-theoretic conclusion from the pieces of the argument that are lying on the table. Let’s try the latter.

3. Arguing for error-theoretic skepticism

Mackie evidently needs to combine proposition (i) with some other premise(s) in order to produce an error-theoretic conclusion. What would serve as a bridging premise is some kind of principle of parsimony that allows the transition from judging that moral facts (or just objective moral facts) have no place in an explanation to concluding that they don’t exist. Both ends of this transition need more scrutiny.

First, showing that moral facts (or objective moral facts) play no role in the explanation of some limited phenomenon, such as disagreement, is surely not going to warrant the wholesale denial of their existence, for (as noted above) moral facts (or objective moral facts) may be needed to explain some other phenomenon. If Mackie is going to have any hope of getting to such a grand anti-realist conclusion, then he must be confident not only that moral facts play no role in explaining moral disagreement, but that they play no role in the explanation of any phenomenon at all. If we are considering *everything* that we know of, then the anti-realist hypothesis is not vulnerable to being overturned by consideration of some other phenomenon that requires a realist explanation, for *ex hypothesi* we know of no such phenomenon.

Second, even it is true that there is no phenomenon whose explanation requires the positing of moral facts, one might still wonder on what grounds someone would prefer denial to agnosticism. The crucial difference is between the following two bridging principles of parsimony:

A) If something plays no explanatory role, then we have no ground for believing in it.

B) If something plays no explanatory role, then we have ground for disbelieving in it.

---

4 The parenthetical asides “(or objective moral facts)” indicate that there are two possible ways of reading Mackie here. We could read him as arguing against the existence of objective moral facts, and then presenting an argument that moral facts are essentially objectivist, thus arguing for a moral error theory. Or we could read him as first establishing that moral facts are essentially objectivist, and then just arguing against the existence of moral facts *simpliciter*. It’s much of a muchness. I will go the latter route, assuming that what is under dispute is the existence of moral facts *simpliciter* (and thus I will henceforth do away with the parenthetical asides).
A philosopher may of course simply announce that (B) is one of his or her basic methodological principles. The logical positivists’ enthusiasm for the verification principle smacked of this kind of programmatic decree (to say nothing of Hume’s rhetorical advice that anything not meeting his empiricist standards should be “committed to the flames”). But simply claiming (B) to be methodological bedrock may seem like a dogmatic overreach when (A) is also available, for surely (A) is more reasonable. Certainly, if we are considering the explanation of a limited phenomenon then this seems to be the case. The fact that unicorns play no role in explaining X (choose any ordinary phenomenon here) may well provide us with no ground for believing in unicorns, but nor does it provide grounds for actively disbelieving in them. But if we lift the limits and consider all known phenomena, then there appears to be more to be said in favor of the stronger principle (B). The fact that unicorns play no role in any good explanation—that, in other words, we have no evidence for their existence whatsoever—may, one might think, be precisely why it is reasonable to disbelieve in them.

Things are somewhat more complicated than this, however. Compare this with a case for which agnosticism intuitively seems the correct epistemic attitude. Are there planets orbiting the star Betelgeuse? There’s currently no evidence one way or the other (so far as I can tell from a quick google search). On the basis of current evidence, one should neither believe that there are planets orbiting Betelgeuse, nor disbelieve this. The presence of planets orbiting Betelgeuse plays no role in explaining any phenomenon we know of, and yet this doesn’t seem to provide grounds for disbelief.

The difference between the two cases is that given the obvious current limitations in our ability to gather evidence about what’s going on in distant solar systems, it is no surprise that we have no evidence one way or the other regarding Betelgeuse’s planets. That we should have any evidence at all about planets orbiting distant stars is still a relatively novel idea; we do not expect to currently have evidence one way or the other regarding a great many stars. By comparison, in the case of unicorns we do expect that if they existed anywhere (on Earth), by this stage we’d have uncovered some evidence of the fact. Perhaps there was a time in the Middle Ages when agnosticism about unicorns was appropriate, but as we explored more and more of the world and uncovered no evidence, the reasonableness of agnosticism gave way to the reasonableness of disbelief. Thus we see that (B) is plausible only with amendment:

B*) If something plays no explanatory role, then we have ground for disbelieving in it, if it is reasonable to assume that if it existed then we would have evidence of it.⁵

So are moral facts more like unicorns or like planets orbiting a distant star?

Answering that moral facts are (in this respect) more like unicorns seems a perfectly coherent thing to say. There is little doubt that this is what Mackie would say, since the claim that if moral

⁵ This amendment should dispel any worries one might have that (B) violates the aphorism “Absence of evidence is not evidence of absence.” Absence of evidence when it is reasonable to expect that presence would leave evidence does count as evidence of absence.
realism were true then we would expect to have evidence of moral facts is very close to the claim he makes in the argument from disagreement: that if moral realism were true, then we would expect to observe fewer moral disagreements. However, it is important to see that plausible options remain open to the realist to resist the skeptical conclusion.

First, of course, moral realists need not accept the claim that moral facts aren’t required to explain anything—they can maintain that moral facts play all sorts of roles in explanations. Nicholas Sturgeon famously considers the case of Selim Woodworth, who in 1846 contributed substantially to the Donner party’s unhappy fate through his ineffective and incompetent leadership of a rescue effort. The historian Bernard deVoto concluded that Woodworth was “no damned good,” and Sturgeon claims that the best explanation of deVoto’s forming this belief is that Woodworth was, in fact, no damned good (Sturgeon 1985). Sturgeon also cites the example of Hitler’s moral depravity as the explanation for his ordering the death of millions of people.

One might object to the realist’s argument by claiming that whenever we have a moral explanation for a phenomenon, there is always a superior non-moral explanation available. After all, we do not need to refer to Woodworth’s being “no damn good” in order to explain his decisions; we could refer instead to psychological factors like his ambition, his lack of empathy, and so forth. In turn, we can explain deVoto’s forming the judgment that Woodworth was no damn good by reference to his having certain beliefs about how Woodworth acted, coupled with deVoto’s commitment to certain moral values which he had come to internalize through a (no doubt complicated) process of socialization. Had deVoto been raised differently, perhaps he would not have condemned Woodworth’s actions in this manner.

But why is the non-moral explanation superior? Why, in fact, is it a competing explanation at all? The anti-realist might try to answer the first question by appeal to parsimony. The non-moral explanation is preferable because it posits less—in particular, it doesn’t require the existence of moral facts. This answer, however, presupposes that the moral facts in question must be something “extra” in an ontological sense, and this is something that a naturalistic moral realist simply denies. The naturalistic moral realist identifies moral properties with naturalistic properties that are already present in the ontological frameworks accepted by all parties involved. If, for example, we can explain Hitler’s actions either by reference to his upbringing, situation, and personality traits (i.e., in non-moral terms) or by reference to his depravity (i.e., in moral terms), then the latter is no more ontologically extravagant than the former if the property of being morally depraved just is the having of those personality traits. By analogy, if we face a choice of explaining a phenomenon (e.g., rust on the exhaust pipe) either by reference to the presence of pairs of hydrogen atoms bonded with single oxygen atoms, or by reference to the presence of water, then neither explanation is ontologically cheaper or costlier than the other. In fact, although one explanation may be more pragmatically suitable than the other to certain conversational contexts, they are not really competing explanations at all.6

---

6 Gilbert Harman (1977, 1986) is sometimes interpreted as claiming that moral facts do not play any explanatory role and therefore we should doubt (or reject) their existence. But he explicitly recognizes that the availability of a naturalistic reduction will save the day for moral facts. And he goes on to suggest particular reductions that he finds
Since we are trying to understand how Mackie might make his argument from disagreement work, it is worth noting that the aforementioned defensive strategy from the moral realist will not move him. Earlier I pointed out that it is a mistake to read the two pages of Mackie’s argument from disagreement in isolation and expect to find a persuasive argument contained therein, for important premises (rejecting noncognitivism, rejecting any non-objectivist construal of moral facts) have already been argued for in preceding pages. (In saying this I am not claiming, of course, that Mackie’s earlier arguments are entirely convincing—far from it—I’m just trying to straighten out the moves.) We find the same pattern here. Prior to offering his argument from disagreement, Mackie has already, to his own satisfaction, deployed considerations against the moral naturalist:

On a naturalist analysis, moral judgements can be practical, but their practicality is wholly relative to desires or possible satisfactions of the person or persons whose actions are to be guided; but moral judgements seem to say more than this. (1977: 33)

Mackie has, in effect, argued that the only avenue available to the moral realist is a Moorean one according to which moral properties really are non-naturalistic ontological “extras.” This is why the obvious realist move of claiming that moral properties do have an explanatory role, in virtue of their being identical to (or supervening on) explanatorily potent non-moral objective properties, won’t cut any ice with Mackie. Because he sees the moral realist as having non-naturalist commitments, he thinks that there are grounds for claiming that whenever we are presented with a moral explanation for a phenomenon, there will always be a superior non-moral explanation available.

The second thing that the moral realist can say against (B*)’s applying to moral facts is that for certain moral facts it is not reasonable to assume that we should have evidence for them (yet). (The corollary of this argument is that it is not reasonable to assume that if there are objective moral facts, then there would be less moral disagreement than there actually is.) Derek Parfit observes that secular ethics is a young discipline, and thus our evidence-gathering methods remain immature (1984: 454). And even when they mature, there may be no guarantee of complete convergence. As David Brink writes: “Moral ties are possible, and considerations, each of which is objectively valuable, may be incommensurable” (Brink 1984: 116; see also Shafer-Landau 1994; Harman & Thomson 1996: 205-6). In such cases, even though there may be moral facts, we would not have reliable evidence one way or the other about them; the evidence would be permanently unclear, and thus disagreement would persist. Some moral facts may simply be enormously difficult to apprehend. While we can be reasonably confident that in principle we could find out whether there are planets orbiting Betelgeuse, we also recognize that doing so currently surpasses our epistemic abilities. Similarly, some moral facts might depend on (for example) a delicate balance of future painful and pleasurable consequences, the knowledge of plausible, ultimately coming to the conclusion that “there is empirical evidence that there are (relational) moral facts” (1977: 132). The moral facts that Harman accepts, though, are not objective in nature, so his view is not a realist one.
which we might suppose that we could *in principle* gain, but for which we recognize that doing so currently surpasses our epistemic abilities. It is worth remembering that realism is, strictly speaking, entirely compatible with radical skepticism, according to which moral knowledge is *impossible*. Of course, the realist is unlikely to be attracted to that position, but it remains an open question what the realist might say about how easily that knowledge comes. Thus the realist might well claim that moral facts are, in this crucial respect, like the planets orbiting a distant star: the fact that we need not posit them to explain any phenomenon doesn’t count against their existence, since it is no surprise that we currently lack evidence.

The discussion thus far has involved a lot of moves and counter-moves, so it may be worth pausing to take stock. We started out wondering how Mackie’s argument from disagreement is supposed to yield his preferred error-theoretic conclusion. The initial question of whether an anti-realist explanation of moral disagreement is superior to a realist explanation of moral disagreement proved to be problematic, since even if the answer were “yes” we would be none the wiser as to whether objective moral facts might be needed to explain some other phenomenon. We were forced to step back and ask the broader question of whether moral facts are required to explain *anything*. Our attention alighted on principle (B), which was then amended to (B*). The issue isn’t whether (B*) is true—let’s assume it is—the issue is whether the principle applies to moral facts. To assess this matter two questions must be scrutinized. The first is whether it is true that moral facts play no explanatory role. I pointed out that Mackie’s positive answer to this question depends on his defeating the possibility of moral naturalism—the arguments for which he presents prior to offering the argument from disagreement. (I have not, however, tried to evaluate Mackie’s argument against moral naturalism, though I happen to think he’s right [see Joyce 2001, 2006: chapter 6; 2016a: 380-1].) The second question is whether it is reasonable to assume that if there are moral facts we would likely have evidence of them. (This is really just a more general way of asking the question posed by the argument from disagreement: if there are moral facts, then shouldn’t there be less moral disagreement than there actually is?) This is a very tricky question to address; it depends very much on the realist’s particular conception of the nature of moral facts.

So we haven’t gotten very far in establishing whether Mackie’s argument from disagreement can be developed into a sound basis for the moral error theory, but we have identified where the battle lines might be drawn, and at least established the perhaps disappointingly exegetical conclusion that while the argument from disagreement gives the illusion of focus—it is, after all, only two pages long—what is really powering it are much larger issues (concerning not just whether moral facts are needed to explain disagreement, but whether they are needed to explain anything) and arguments that lie elsewhere.

---

7 David Enoch writes: “A radically inaccessible realm of moral facts is, I think, a very small comfort for the realist. Such realism may, at most, serve as a last resort, but it is to be avoided if at all possible” (2009: 22). And Brink writes that if the realist were to rely too much on appealing to ties and incommensurability, it “would weaken his reply to the argument from disagreement” (1984: 116).
4. Arguing for justification skepticism via genealogical debunking

Rather than wondering how Mackie might press this argument for error theory more effectively, now I want to consider the possibility of his backing off from that strong conclusion and aiming for a less ambitious form of skepticism (not all-things-considered, but just with respect to the argument from disagreement). Suppose he were to accept (A) instead of (B*):

A) If something plays no explanatory role, then we have no ground for believing in it.

Because (A) is more modest than (B), there is no need to amend it in the analogous way that (B) became (B*). And because these are bridging principles, the conclusion at the far end of the bridge can be proportionally weaker. We are no longer aiming for (ii) or (iii), but their less presumptuous counterparts:

ii*) … and there may be no objective moral facts.

iii*) … and there may be no moral facts.

The anti-realist still has to establish that moral facts (or objective moral facts) play no explanatory role, and we know that the realist need not concede that point without a serious fight. Most of what I have already said about that debate holds as much for this weaker argument as the previous stronger one. But the anti-realist has one less task to do: he or she doesn’t have to argue for the difficult claim that it is reasonable to believe that if objective moral facts existed, then we would have evidence of them. Recall again the Betelgeuse case: I noted that the presence of planets orbiting the star plays no role in explaining any phenomenon we know of, and yet this doesn’t seem to provide grounds for disbelief, since currently we wouldn’t really expect to have evidence one way or the other. However, it’s still reasonable to declare that because the presence of planets orbiting Betelgeuse plays no role in explaining any phenomenon we know of, for all we know there may be no such planets. In other words, though we might not have grounds for disbelief, nor do we have grounds for belief.

Such a position still counts as skeptical, though it is not the skepticism of the error theorist. Rather, just as we would say that someone who currently believes that there are planets orbiting Betelgeuse lacks justification for this belief and therefore lacks knowledge, so too (according to the view under discussion) someone who has any moral belief lacks justification and therefore lacks knowledge. Calling this conclusion “weaker” than the error-theoretic result shouldn’t lull us into failing to notice that it is still a radically skeptical view.

The argument under consideration is essentially a genealogical debunking one. Start again with hypothesis (i) (moral codes “reflect people’s adherence to and participation in different ways of life’”—an explanation that appears not to require the existence of moral facts. Again: the question of how to explain moral disagreement is just the point of departure; what really matters
is whether moral facts are needed to explain anything. But as soon as we make the issue broader in this way, we should see that we can immediately shrink it again to the question of whether moral facts are needed to explain moral judgments. Let me explain. There is no phenomenon that we might be tempted to explain by reference to moral facts without our having made a moral judgment being an indispensable element. Consider Sturgeon’s example of Hitler’s actions, for instance, which he wants to explain by reference to Hitler’s depravity. But one will be tempted by that explanation only if one is willing to judge Hitler’s character depraved in the first place, and then the question becomes what explains that judgment: must we invoke Hitler’s depravity to explain why someone judges him depraved, or can that judgment be explained better by a hypothesis that refers only to non-moral phenomena? If any person’s moral judgment about Hitler can always be best explained without reference to moral facts, then we might conclude that those judgments lack justification. And if one’s judgment that Hitler was depraved lacks justification, then no explanation that appeals to his depravity as an explanans (to explain genocide, etc.) should be accepted. In other words, the question of whether moral facts need be invoked to explain any phenomenon always boils down to the question of whether they need be invoked to explain moral judgment.

There has been quite a bit of discussion in recent years about genealogical debunking arguments (see Joyce 2006, 2016b, 2016c; Wielenberg 2010; Kahane 2011; Fraser 2014; Braddock 2016). The discussion has often taken the evolutionary perspective: arguing whether human moral thinking is the product of natural selection and, if so, what its adaptive purpose might have been. The evolutionary debunking argument has promise because it’s reasonably plausible both to claim that moral thinking is the product of natural selection, and to claim that its evolutionary purpose was to play a role in strengthening our ancestors’ social bonds so as to encourage them to cooperate together more effectively. The significance of the second claim is that moral thinking is explained in a way that makes no reference to any moral judgments being true. In this it contrasts with evolutionary accounts that might be given of other pieces of human psychology, such as our ability to recognize faces, for example. The evidence seems to indicate that humans have an innate mechanism for visually distinguishing faces from other stimuli (see Slater & Quinn 2001). In explaining why such a mechanism might have evolved, one is likely to mention the importance of social bonding early in infancy, the stability of the presence and anatomical structure of human faces, and so on. The crucial feature of this explanation, though, is that it presupposes that faces actually existed in the ancestral environment—the face-identifying mechanism was useful precisely because it reliably succeeded in putting the infant in causal contact with actual faces. The evolutionary account of the human capacity to make moral judgments differs critically in this respect. The explanation is (very roughly) that having the ability to judge certain actions as morally required (say) was useful because it encouraged our ancestors to cooperate in fitness-enhancing ways—not because it allowed them to identify which actions really did have the property of being morally required. The view is not that particular moral judgments are hard-wired; it is that the basic capacity to employ a moral conceptual framework is hard-wired, and then the social environment determines which moral norms one
ends up endorsing. (In an analogous way, humans may have evolved mechanisms dedicated to acquiring a language, but precisely which language a person ends up speaking is determined by the social environment.)

Despite the focus on the evolutionary perspective in recent literature, a genealogical debunking argument can run without it. What matters is that a plausible, or perhaps even empirically confirmed, complete account of moral judgment can be given which does not imply or presuppose that moral judgments are true or even probably true. Instead of the explanation being at the evolutionary level, it might instead be at the anthropological level (how cultures come to adopt their moral norms) or at the level of developmental psychology (how individuals come to internalize moral norms).

But any genealogical argument is susceptible to the same objection that we encountered earlier: even if moral judgments can be explained entirely in non-moral terms, this does not show that moral facts are explanatorily impotent if moral facts are identical to those non-moral facts mentioned in the explanation. Forms of moral naturalism promise to establish just such an identity relation.8

One strategy for the proponent of the debunking argument, then, would be to supplement it with anti-naturalist arguments. (These arguments would just be against moral naturalism, of course; they could be entirely consistent with a more general commitment to methodological naturalism.) This is the strategy I pursued on an earlier occasion (Joyce 2006), peddling arguments that can be seen as elaborations of the same doubts that Mackie voiced against moral naturalism, quoted briefly earlier. In fact, however, such arguments might be considered overkill. All that really needs to be established is that particular versions of moral naturalism are unacceptable—namely, those that would relate moral facts to those naturalistic properties mentioned in the genealogical explanation. Thus the proponent of the debunking argument need not have a prior commitment opposed to the very idea of moral naturalism.9

Another strategy for the proponent of the debunking argument (and, for what it’s worth, one that I’ve preferred in recent years) is to see the argument in terms of establishing a burden of proof. It’s not enough to say that there might be identity relations holding between moral properties and those naturalistic properties explicitly cited in the genealogical explanation—rather, the proof of the pudding is in the eating. After all, even the moral error theorist is going to agree that moral judgments have some kind of history—whether at a psychological, anthropological, or evolutionary level—and someone can always claim that it’s possible that the moral facts are surreptitiously buried somewhere among the historical facts that the error theorist is willing to accept. However, it seems reasonable for the error theorist to remain unbothered by this claim until the naturalistic account is displayed and defended. The conclusion is not that we

---

8 My focus on the identity relation is largely for brevity. In fact, other weaker nomological relations, such as supervenience or some kind of probabilistic causal relation will suffice to undermine the debunking argument (see Brosnan 2011: 61). The difference is very important in other areas, but not, I think, to any of the arguments discussed in this chapter.

9 Contra Das 2016.
should provisionally assume that the moral error theorist is correct in thinking that moral judgments are false; the conclusion is that since moral judgments are the product of a process that appears to be consistent with a moral error theory, then we should provisionally assume that they lack justification.

Consider an analogy. Suppose it turns out that when asked about the capital of Gabon I find myself spontaneously answering “Freetown.” I assume it’s something I picked up somewhere, though can’t recall any details. But now you show me some credible video footage from last Monday of someone asking me the capital of Gabon while I shrug and answer that I have no idea. Mysterious. Then you show me credible footage from last Tuesday of my being hypnotized by Madame K., who is telling me “When you awake, you shall believe that the capital of Gabon is Freetown, and you shall forget all about being hypnotized.” Mystery solved. But now what attitude should I take toward the proposition that Freetown is the capital of Gabon? I could just look it up, but let’s say that for some reason I haven’t yet had the opportunity. It’s clear that I really don’t know what to think anymore. Maybe Freetown is the capital of Gabon, maybe it isn’t. It would be reasonable to conclude that my previous fairly confident belief lacked justification (or, at least, that if I persisted with the belief it would now lack justification). It lacks justification because I’ve discovered that it’s the result of a process that is consistent with its not being the case that Freetown is the capital of Gabon and does not render that fact very probable.

Perhaps I ask my friend Mary about the capital of Gabon, and am relieved to hear her confidently claim that it is Freetown. Justification reinstated! But then Mary and I are shown footage of her also being hypnotized by Madame K. last week (“The capital of Gabon is Freetown … the capital of Gabon is Freetown”), so now Mary and I are in the same confused epistemic state. And so we should be, for we have learned that our belief has come about through a process that disconnects it from the relevant facts. Mary has an idea: “Perhaps Madame K. hypnotizes people to believe only true things about countries’ capitals!” This idea amounts to suggesting that the process that produced our belief does connect to the relevant facts (concerning African countries and cities), though somewhat more indirectly than we’d previously thought. But the mere possibility of this connection—of Madame K.’s being epistemically benevolent in this manner—is not sufficient to reinstate justification. For that, we’d need some credible evidence that Madame K. actually does behave in this manner. In the same way, the mere possibility of a naturalistic theory connecting moral facts to the non-moral facts that figure in the error theorist’s genealogy of moral judgment is not sufficient to reinstate justification. For that, the naturalistic theory needs to be made credible.

Whichever strategy the debunking skeptic prefers—whether going on the offensive and trying to refute versions of moral naturalism (or moral naturalism simpliciter), or defensively claiming that it is up to the naturalist to put forward a credible theory—it is clear that the debunking argument isn’t designed to be an argument that defeats moral naturalism, but rather one that requires supplementation with anti-naturalist considerations. Recall that the same thing goes for the stronger error-theoretical skeptical argument that I examined earlier—the one revolving around (B*). This argument would also fail if certain versions of moral naturalism
could be made plausible. It is also clear that these complex skeptical arguments have brought us a long way from the original presentation of Mackie’s argument from disagreement.\(^\text{10}\)

5. Arguing for justification skepticism via disagreement among peers

There is a more direct route from moral disagreement to justification skepticism. Proponents of the so-called \textit{conciliatory view} on the epistemic significance of disagreement hold that when one encounters a disagreeing epistemic peer, one’s epistemic confidence in the disputed claim should diminish (see Feldman 2006; Christensen 2007; Kornblith 2010; Matheson 2015).\(^\text{11}\) The basic idea is quite mundane: suppose I glance at my watch and it reads 11:15; I therefore confidently believe that the time is 11:15; but then I notice that the clock on the wall reads 11:35. If I lack any grounds for privileging my watch over the clock, the fact that the clock “disagrees” with my watch should immediately give me pause. The confidence that I had in the belief that the time is 11:15 should be reduced; if I care about knowing the time, then I’ll need to take some steps to sort out the discrepancy. In an analogous manner, if you and I are splitting the tip at a restaurant, and I come to the result of paying $11.15 each but then you come to the result of $11.35 each, and I am as confident in your arithmetical abilities and honesty as I am in my own, then we have a puzzle: I should reduce my initial confidence in $11.15 as the answer and recalculate the sum more carefully. In an analogous manner again (supposedly), if I’m confident in judging that x is morally wrong, but I then encounter someone who thinks that x is morally acceptable, and I have no ground for privileging my own moral judgment-formation processes over those of the dissenter (i.e., I must accept that he or she is an epistemic peer in this matter), then my confidence in judging that x is morally wrong should be reduced.

Some have thought that disagreement with epistemic peers is so ubiquitous that a completely global skepticism follows. This was, famously, an important kind of argument put forward by classical skeptics of the Greek and Roman philosophy worlds.\(^\text{12}\) Here we are interested in a more limited but still remarkable result: that there is something about moral disagreement in particular that leads, via a version of this argument, to moral skepticism. I have already remarked (regarding the argument from quality of disagreement) that moral disagreement seems to have qualities that are less characteristic of many non-moral disputes (intractability, etc.); perhaps this provides the basis of rendering moral skepticism plausible while allowing us to avoid the extravagant pessimism of global skepticism. As quick evidence of the difference, consider again

\(^{10}\) Tersman argues that consideration of moral disagreement may re-enter the debunking debate later: as a factor potentially counting against the moral naturalist’s ability to defeat the debunking argument by providing a theory that plausibly connects moral facts to the non-moral facts accepted in the debunker’s genealogy. See Tersman 2015b.

\(^{11}\) Others reject the conciliatory view in favor of the \textit{steadfast view}: that it is acceptable to remain confident in one’s beliefs in the face of disagreement from epistemic peers (see Kelly 2010; Sosa 2010; Lackey 2010). Here I am accepting a conciliatory view for the sake of argument.

\(^{12}\) Though Diego Machuca (2015: 27) has rightly argued that it is a mistake to read too much of the detail of the modern debate about the epistemological significance of disagreement into the views of the ancient skeptics.
the case of you and me coming up with different numbers when trying to split the restaurant tip. As mentioned, we would probably proceed by recalculating the sum more carefully. But suppose we both do so and I again come up with $11.15 and you again come up with $11.35. We frown and try again and the same thing happens. At this point we’d just be utterly baffled; it’s not clear what we should do. But in the case of moral disagreement we’re reasonably tolerant of the possibility that no matter how much you and I deliberate carefully and “compare notes,” I may simply continue to find x morally wrong while you continue to find it morally acceptable. We’re not surprised that moral disagreements can persist in this manner; it’s not baffling.

Central to the conciliatory view is the idea that one must be able to identify epistemic peers in a reasonable manner: those whose intelligence, freedom from bias, reflective awareness, access to and appreciation of the evidence, etc., are equal to one’s own. One must, in short, be able to form justified views about others’ epistemic credentials. Most proponents of the conciliatory view add some version of the independence principle: that in identifying epistemic peers one must discount the fact of the dispute in question—you cannot, in other words, take the very fact that the person disagrees with you as evidence of her epistemic inferiority. If the dispute is over whether \( p \) is the case, then in evaluating your opponent’s epistemic credentials you must ignore your belief that \( p \) and the reasoning that led you to that belief (and ignore her disbelief that \( p \) and the reasoning that led her to that disbelief). This principle promises to block the slide to wholesale skepticism. If a person disagrees with you about something incredibly fundamental, like whether the material world exists, then if, in assessing whether he is an epistemic peer, you must discount this belief of his and all the reasoning that led him to it, then it is unlikely that you will be left with sufficient resources to make a judgment of his epistemic credentials, in which case conciliationism simply remains silent on whether you should revise your belief on the matter (see Elga 2007; Vavova 2014).

Let us assume for the sake of argument that the conciliatory view is basically correct. Would moral skepticism follow? I will restrict myself to commenting on two reasons for thinking that it would not.

First, how one goes about assessing whether a moral disputant is an epistemic peer seems dependent on one’s standing general attitude to the status of moral facts and evidence. Suppose Mary has already taken on board some of the worries that have been canvassed earlier in this chapter: that moral facts do not seem to play a role in explaining any phenomenon, that it’s not clear what would count as evidence for the truth of one moral judgment over another, that moral judgments can be explained in a way that appears consistent with an error theory, etc. Despite harboring these worries, Mary hasn’t endorsed moral skepticism and continues to make moral judgments: she judges that x is morally wrong. She encounters Fred, who disagrees. Mary must now assess whether Fred is her epistemic peer. She can presumably wonder whether he is an epistemic peer on non-moral matters, but what about on moral matters? (It would seem strange to think that all she need concern herself with is whether Fred is an epistemic peer on non-moral matters. That would be like claiming that when Mary wonders whether Fred’s disagreement over a math problem undermines her belief on that point, it suffices for her to ascertain whether he is
her epistemic peer concerning horticulture.) While Mary can consider whether Fred is as intelligent and reflective as she is, how shall she assess whether Fred has equal access to and appreciation of the moral evidence? The natural concerns that non-moral disagreement might prompt—such as “Maybe he has better access to the evidence than I do” or “Maybe he appreciates the evidence better than I do”—may seem to Mary to be not even sensible worries to voice about moral beliefs. She’s just not sure, in other words, whether, for moral cases, it makes sense to ask whether Fred is an epistemic peer. And if she cannot assess whether Fred is an epistemic peer, then the conciliatory argument remains silent on whether she should alter her confidence in the moral claim that x is wrong.

What is interesting in these thoughts is that Mary is sympathetic to views that (as we have seen earlier in this chapter) can be used as premises in arguments for moral skepticism, but here those very same views are blocking the skeptical result. And the role of moral naturalism in the argument is reversed as well. In earlier arguments, a viable version of moral naturalism promised to thwart the skeptical argument. It is therefore somewhat ironic that here if a viable version of moral naturalism were forthcoming, then Mary’s worries would be answered, and she would be able to assess Fred as an epistemic peer (ceteris paribus), in which case the possibility that Fred’s disagreement should lead Mary to downgrade her confidence in her moral judgment (i.e., the moral skeptic-friendly possibility) would reopen.

A second reason for thinking that moral skepticism does not follow from conciliationism follows from the “discounting” of beliefs required by the independence principle. Moral beliefs, the thought goes, are unlikely to be held in isolation from each other. If Mary and Fred’s disagreement about whether x is morally wrong is quite fundamental, and if, in assessing whether Fred is an epistemic peer, Mary must discount her belief that x is morally wrong and all the reasoning and evidence that led her to that belief, then her resources for deciding whether Fred is an epistemic peer diminish. Katia Vavova pushes this argument, maintaining that “as our disagreement deepens, the grounds I have for taking you to be my peer shrink” (2014: 314). If, on the other hand, Mary and Fred’s moral disagreement is relatively shallow—if, that is, they have a track record of agreeing on most moral matters, but have encountered this one difference in moral opinion—then an epistemic retreat from confidence to agnosticism, as conciliationism appears to demand, would be entirely appropriate. But since (Vavova thinks) such appropriate reductions in moral confidence in response to disagreement are sufficiently rare and reasonable, then they represent no comment-worthy concession to moral skepticism.

It seems to me that this argument is a hostage to empirical fortune. It is far from clear to what extent people’s moral beliefs are based on “reason and evidence,” and the assumption that people derive their moral judgments from broader moral principles, which in turn are based on the endorsement of fundamental moral values, may be something of a moral philosopher’s optimistic projection of Ethics 101 onto human moral psychology. While it seems true that moral judgments often come in “packages” (e.g., in the US, someone’s views on abortion likely correlate with their views on gun control and taxation), it doesn’t follow that this is because these views have been inferred from deeper moral principles. It may be, rather, that certain personality
traits, like risk-aversion and disgust-sensitivity, lie behind these judgments (see Choma et al. 2014; Hibbing et al. 2014), and that lying behind some of these differences in personality traits is genetic variation (see Benjamin et al. 2012; Hatemi & McDermott 2012). I am simplifying things exceedingly in order to make the point succinctly—the point being that there may be causal connections among packages of moral judgments (e.g., being jointly caused by some neurological trait), but the demand made by the independence principle—that one must ignore the disputed belief and all the reasons that have led one to it—concerns the inferential process that has led one to the belief, not the belief’s psychological causal antecedents.

Vavova’s view is that in deep moral disputes the independence principle requires one to bracket off so much that one no longer has grounds to judge whether one’s disputant is an epistemic peer. I have raised a doubt about this, since it seems to assume that human moral psychology works in a manner that it quite possibly does not work. But this represents no great victory for the moral skeptic. Perhaps in cases of deep moral dispute one is not required to bracket off so much that one cannot judge whether one’s disputant is an epistemic peer, but if the disputant nevertheless disagrees with a great many of one’s moral beliefs, then one may simply conclude that the disputant is not an epistemic peer. The difference is between “I can’t tell whether you’re an epistemic peer” and “You’re not an epistemic peer.” Either way, one lacks ground for taking the person to be an epistemic peer, and so conciliationism doesn’t require a downgrade in the epistemic status of one’s belief on the disputed matter.

But I think this continues to make the avoidance of moral skepticism a hostage to empirical fortune. It cannot simply be assumed that whenever someone disagrees with me on a certain weighty moral matter, this person must also disagree extensively on other matters so much so that I will deny that he or she is an epistemic peer. It’s an empirical question to what extent human moral judgments can be compartmentalized. It might be responded that if Mary has reason to suspect that Fred’s moral disagreement with her is due to some odd psychological compartmentalization on his part, then this disqualifies him from being considered an epistemic peer. But this seems contrary to the spirit of the epistemic modesty that conciliationism champions. Mary might well suspect that Fred’s moral judgments are due to quirky aspects of human psychology rather than consistent inferences from more basic values (perhaps Fred’s view on this one matter has been influenced by arbitrary factors), but epistemic modesty should leave her wondering whether her own heartfelt moral judgments are any better off. Mary may come to the conclusion that Fred is epistemically flawed in his moral judgments, but if she wonders whether she too is flawed in similar ways—that this is just how human moral psychology works—then this is no reason to deny that Fred is her peer.

It is also worth noting here that the kind of people who might disagree with a weighty and seemingly obvious moral truth do not need to be construed as moral monsters (Vavova mentions Caligula and Clarisse the “homicidal sociopath” [2014: 314])—they may instead be friendly and bespectacled metaethicists who happen to be error theorists about morality, or maintain skeptical views about certain elements of morality. In ordinary contexts, if someone disagrees with your claim that it is morally wrong to shoot strangers, then you’d take him to mean that it’s morally
permissible to shoot strangers, in which case you’d be appalled, think there was something wrong with him, and probably downgrade his epistemic credentials. But in the context of doing metaethics, this inference would be mistaken: an error theorist may deny that it is morally wrong to shoot strangers, but she should quickly append that nor is it morally permissible to shoot strangers—it’s not morally anything—and she might add that she is adamantly and passionately opposed (on non-moral grounds) to shooting strangers. (Gilbert Harman once claimed on metaethical grounds that we have no business saying that it was wrong of Hitler to have ordered the extermination of the Jews, though he quickly appended that we can say that Hitler was an evil man and that his actions brought about something that ought never to have happened [1975: 7].) Well, I suppose that one might still feel appalled at this, but it’s a quite different sort of appallingness than the former kind, and it’s considerably less obvious that espousing such a view should count immediately against one’s epistemic credentials.

Conciliationism offers a more direct route from moral disagreement to justification skepticism than genealogical debunking arguments, but the route is far from trouble-free. One thing I’ve stressed is that how one assesses another’s epistemic credentials may differ in moral cases from non-moral cases, due to the psychological mechanisms lying behind moral judgment being substantially different from those that lie behind other kinds of judgment; and an interesting take-home message is that many of the relevant details remain unknown empirical territory.

6. Conclusion

There are different kinds of moral skepticism, and various possible paths to each of these views. Many of those paths have nothing to do with the phenomenon of disagreement, and among these may be the strongest skeptical arguments. This chapter has focused on several entwined arguments that do focus on moral disagreement, endeavoring to display their interrelations and difficulties.

References


