Knowing What Matters[*]

Richard Yetter Chappell[†]
Princeton University

January 14, 2012

Abstract

Parfit's *On What Matters* offers a rousing defence of non-naturalist
normative realism against pressing metaphysical and epistemological
objections. He addresses skeptical arguments based on (i) the causal
origins of our normative beliefs, and (ii) the appearance of pervasive
moral disagreement. In both cases, he concedes the first step to the
skeptic, but draws a subsequent distinction with which he hopes to
stem the skeptic's advance. I argue, however, that these distinctions
cannot bear the weight that Parfit places on them. A successful moral
epistemology must take a harder line with the skeptic, insisting that
moral knowledge can be had by those with the right kind of psy-
chology — no matter the evolutionary *origin* of the psychology, nor
whether we can *demonstrate* its reliability over the alternatives.

1   Causal Origins-based Skepticism

1.1   The Naturalist Argument for Normative Skepticism

There are any number of logically consistent moral worldviews that a person might endorse. Many of us hold that pain is always and everywhere bad. But we can just as well imagine aliens who consider pain and suffering to be intrinsically good,[1] or "Future Tuesday Indifferent"[2] agents who hold that pain only matters when experienced on days other than Tuesday. Such values strike us as bizarre, but we can nonetheless imagine agents who hold them sincerely, and who consider *us* bizarre for denying what seems to them 'obvious' facts about the value of pain. Call this the fact of *coherent moral diversity*.

According to Parfit's non-naturalist normative realism (hereafter, 'moral realism'), there's an objective fact of the matter about the dis/value of pain. There's one correct view, and all other possible views are — to a greater or lesser extent — mistaken. This then raises an obvious epistemological challenge for realists who also take a stand in favour of one first-order moral viewpoint over others: Why think that *their* moral view, amongst all the coherent possibilities, is the correct one?

The problem is especially acute for those of us who share Parfit's view that the moral facts, being non-natural, are causally inert. For this means

---

[1] I will continue to speak of "pro-pain" agents, for vividness, but if you doubt the coherence of this particular example, feel free to substitute a less extreme case of moral error.

[2] Parfit (1984, 124)

that we cannot consider our moral beliefs to be caused by the moral facts, the way we think that our perceptions of physical objects are typically caused by those objects. Such a causal connection could explain the reliability of our beliefs. In its absence, we seem to be left without any grounds for taking our moral belief-forming mechanisms to be at all reliable. After all, there is going to be some purely natural, causal explanation of why we've ended up with the particular moral beliefs that we have — perhaps appealing to evolutionary, psychological, or sociocultural factors — a causal explanation that at no point invokes the moral facts themselves as playing any role in shaping our moral intuitions. Call this *the causal irrelevance of the moral facts.*

Critics like Sharon Street then ask: Shouldn't we then think it extraordinarily *unlikely* that those natural causes would happen to lead us to the one true morality? Given the fact of coherent moral diversity, plus the causal irrelevance of the moral facts, realists seem forced to admit that it would be a "striking coincidence" (Street ms) if our moral beliefs turned out to be (anywhere near) correct. We shouldn't believe something so improbable. So if we accept moral realism, we're committed to thinking ourselves totally ignorant and incapable of grasping the moral truth. This argument is a version of what Parfit calls "the Naturalist Argument for Normative Skepticism" (*OWM*, II, 513).

1.2 Parfit's Response

First consider how we could respond if the argument's target (whose reliability was in question) was not ourselves, but some alien species. In the absence of any more particular information, we have no *antecedent* reason for expecting some arbitrarily chosen alien species to be morally reliable. But rather than trying to assess their reliability more or less a priori, with little or no concrete information, we would do better to inquire into their concrete circumstances. If we look and find that the aliens seem generally sympathetic and altruistic, concerned to promote the wellbeing and non-harmful life goals of other sentient beings, then we may be reasonably confident that they're on the (morally) right track, whatever the causal process that brought this about. If, on the other hand, they seem to enjoy gratuitous torture, then we'll judge them to be morally abhorrent.

Moving up a level of abstraction, we may also have some idea of what evolutionary and socio-cultural pressures are most likely to bring about these morally good practices in an intelligent species. For example, we may expect sympathy and altruistic norms to more likely arise in social animals where there are great gains to cooperation, and where pair-bonding and parental investment are important for long-term reproductive success. The precise details don't matter for my purposes. The point is just that, once we learn that a species lives in a certain kind of environment, it will no longer be a "coincidence" if they turn out to have generally correct views about the value

of friendship, love, sympathy, aversion to pain, etc.[3] And so it is for human beings: We (as observers, using our best moral judgment) may reasonably conclude that we (as anthropological subjects) are, for all our manifest faults, generally pretty reliable on moral matters!

Of course it sounds hopelessly circular to use our own moral judgments to assess our moral reliability like this, and it certainly won't persuade the skeptic — but trying to persuade the committed skeptic is a fool's game. It's logically impossible to give a non-question-begging justification to one who questions *everything*. So that cannot be our goal. A more appropriately modest philosophical goal is simply to provide an *internal* defense of our claims to knowledge, showing how the realist — given her starting assumptions — could reasonably fail to be swayed by the skeptic's argument. The question for us is thus not whether the skeptic must, on pain of inconsistency, grant the realist's claim to moral knowledge (of course he needn't), but whether the realist is forced to accept the skeptical conclusion that she lacks moral knowledge. We have nowhere else to start but from the premises that seem to us to be true, and from this starting point there's nothing obviously self-undermining about the moral realist's worldview.

In his response to the Naturalist Argument for Normative Skepticism, Parfit similarly defends the need for making normative assumptions along the way:

> Some whimsical despot might require us to show that some clock

---

[3] Cf. Copp (2008)'s 'quasi-tracking thesis', and Enoch (2010).

is telling the correct time, without making any assumptions about the correct time. Though we couldn't meet this requirement, that wouldn't show that this clock is not telling the correct time. In the same way, we couldn't possibly show that natural selection had led us to form some true normative beliefs without making any assumptions about which normative beliefs are true. This fact does not count against the view that these normative beliefs are true. (*OWM*, II, 533)

After this point, however, Parfit takes an odd turn. He concedes to the skeptic that if our "normative beliefs were mostly produced by evolutionary forces... that would count strongly against the view that we can respond to [their] intrinsic credibility" (534). He then spends several pages arguing that our normative beliefs are not plausibly explained by evolution.

This is puzzling, for two reasons. Firstly, if evolutionary forces didn't cause our normative beliefs, that just means that some other natural cause did. This makes Parfit's fixation on specifically *evolutionary* causes baffling. Why should evolutionary forces be seen with any more or less suspicion than any other (equally non-moral) natural cause?

Secondly, the quoted passage makes it sound as though "evolution" and "intrinsic credibility" are two rival causal explanations for our normative beliefs. But Parfit acknowledges that intrinsic credibility is not causally efficacious (*OWM*, II, 502). Presumably the way that we respond to the intrinsic credibility of a proposition is by being psychologically constituted

such that we are disposed to believe the right things (*OWM*, II, 503; see also §2.1, below). Evolutionary forces may well be a large part of the causal story of how we got to be so constituted. And if not, there will be some other, similarly natural, causal story. Either way, there's no necessary tension between having the right psychology and having an independent (non-moral) causal explanation for how we came to possess this psychology.

Parfit seems to recognize this in the case of evolution via group selection: "When the acceptance of certain normative beliefs made some community or culture more likely to survive and flourish, this fact does not *as such* cast doubt on the truth or plausibility of these beliefs." (*OWM*, II, 537, emphasis added.) But surely the point generalizes: it's just as unclear why other causal origins should *as such* be epistemically undermining, independently of the substantive content of the beliefs thereby produced. Some special explanation must be offered for treating different kinds of causes differently, given that they are alike in the salient respect of being natural rather than normative.

Perhaps Parfit is assuming that individual-level selection is more likely to yield selfish rather than pro-social norms. That is, holding fixed our understanding of what particular things are right and wrong, one might expect individual-level selection to cause creatures to endorse the latter rather than the former. But then it is really the (presumed) *content* of these norms, rather than their origins, that Parfit is objecting to. This would render unnecessary his argument that our pro-social norms aren't explained by evo-

lution. So long as they have the right content, their origin shouldn't matter. Since we actually endorse pro-social norms, we know (trivially) that evolution did *not* lead us to endorse the contrary norms that we take to be wrong, so we have no content-based reason for taking our beliefs' possible evolutionary origins to be undermining. If evolution caused our actual beliefs, then the beliefs it caused are pro-social, for those are the beliefs that we actually have.

Parfit thus faces a dilemma. What he really needs is a purely *formal* reason for thinking that (individual-level) evolution, more than other natural causes, would likely have led us morally astray. But absent our substantive normative assumptions — from the perspective of a moral 'blank slate' — there would be no basis for any such judgment. On the other hand, given our *substantive* normative assumptions (and the psychological fact that those are our normative assumptions), such a claim is trivially unfounded.

I wish to propose a more flat-footed response on behalf of the realist. Rather than relying on this dubious distinction between evolutionary and other causes, and holding our moral epistemology hostage to the contingencies of an unsettled empirical question, realists should cut off the skeptic's argument at its root — namely, its assumption that the causal origins of our normative beliefs are ever *in themselves* epistemically undermining.[4] The basic response to the skeptic's argument is then simple: All things consid-

---

[4] The italicized proviso is necessary to rule out trivial cases where, e.g., one is caused to believe a moral claim previously known to be false. In that case it is not the cause per se, but more fundamentally the resulting belief, that is the problem.

ered, in light of all that we know or believe to be true, we have every internal reason to retain our "default trust" in our normative beliefs.[5] From our current standpoint, we have every reason to regard our pro-social evolutionary heritage as providing us with roughly correct moral intuitions (along with various biases which we can hope to identify and dispel by reasoning from our core moral commitments). So reflection on the causal origins of our moral beliefs is not in itself epistemically undermining. It would only be undermining if we accepted the skeptical principle that we need to provide an *independent* justification for our beliefs — a justification that would convince even one who has adopted a position of radical doubt towards the beliefs in question. But we should not accept such a skeptical principle. So it's not clear that there's any real force to the Naturalist Argument for Normative Skepticism after all.

1.3   The Moral Lottery

Street (ms, 26) acknowledges that "in seeking to answer the skeptical challenge...the normative realist is...entitled to offer an answer that is ultimately question-begging." However, she goes on to claim that the flat-footed realist reply "provides no *reason* to think that the causal forces described by our best scientific explanations shaped our normative judgments in ways that might have led those judgments to track the truth; it merely confidently reasserts that they did." Here the dialectic becomes murky: Given our prior

---

[5] On the importance of "default trust", see Railton (2004), and §2.2, below.

belief that (e.g.) helping others is good, why does this not qualify as an internal reason to think that pro-social evolutionary forces put a species on the morally right track? It's not an independent or non-question-begging reason, granted, but that doesn't yet show that it is no reason at all.

Street's subsequent discussion of perceptual epistemology suggests the following view. We may seek to explain the accuracy of our mental faculties in either of two ways. Firstly, we might presuppose some particular, substantive claims in the domain in question, and then merely show how our mental faculties would lead us to believe those putative truths. This is the kind of explanation — call it *substantive explanation* — that the moral realist offers. By contrast, a *constitutive explanation* aims to show how our mental faculties can be reliable merely in virtue of general facts about the nature of the domain in question. For example, it is in the nature of concrete objects to have causal powers, which offers a general explanation of how we might reliably detect them. Similarly, on Street's constructivist anti-realism, it is in the nature of normative facts that they be coherent outgrowths of our core practical commitments, which suggests a general explanation of how reasoning from our core practical commitments might be expected to lead us to true moral beliefs. On Street's view, these constitutive explanations are epistemically legitimate (though the radical skeptic could deny even these), but substantive explanations are not. Call this *Street's principle.* My claim is then that moral realists should simply reject Street's principle. When engaging in wide reflective equilibrium we may appeal to *all* our beliefs,

including our particular substantive beliefs within a domain, as providing (defeasible) internal reasons.

But even if the most general form of the skeptical argument thus fails, perhaps there are ways of expanding upon it to show that the realist is (by their own lights) in epistemic trouble. Street's lottery analogy suggests one such route:

> [The realist's response] is no better than insisting, without any non-trivially-question-begging reason to think so, that one has won the New York Lottery. Given the odds we can reasonably suppose to be in play in this "normative lottery" case, we should conclude that in all probability we *didn't* win—that, if there is indeed such a thing as the robustly independent normative truth we are positing as a substantive normative premise, then we are probably among the unlucky ones who (just like the ideally coherent Caligula, grass-counter, hand-clasper, and so on) are hopeless at recognizing it. (21)

To restate the argument: We all accept epistemic principles that disallow believing (a priori) oneself to have won the lottery against all odds. And, Street assumes, the fact of coherent moral diversity plus the causal irrelevance of the moral facts together establish that the non-skeptical moral realist is in an epistemically equivalent position of believing herself to have won the "normative lottery" against all odds. So those same epistemic principles

must similarly disallow the realist from complacently believing herself to be so lucky.

The main problem with this analogy concerns the "odds" that we assign to the various outcomes. Given our understanding of the physical lottery mechanism, we know that we should assign equal odds to each possible outcome, and hence the odds we should assign to some particular ticket (#139583923, say) winning is extremely low. For Street's argument to work, it must be that we are similarly required to assign equal (or roughly equal) odds to the truth of each possible normative system. But *why think that*? I certainly have no antecedent inclination to assign equal odds to all possible normative systems — I think it's overwhelmingly more likely that pain is intrinsically bad than that it's intrinsically good, and I don't see anything in Street's argument that suggests I should change my mind. (It's not as though the normative truth was itself settled by a chance process that gave equal objective probability to each possible outcome.)

Street might try to push the analogy by asking us to imagine a lottery player who likewise insists that the (a priori) odds of ticket #139583923 winning are much higher than for any other ticket. There are a couple of things to say about this. Firstly, I agree that such a person would be crazy. But my reasons for thinking this depend on a fact specific to the physical lottery case, namely, the role of a chance mechanism in producing the outcome. A person who fails to assign equal odds across the possible outcomes of a fair lottery either doesn't understand the underlying physical mechanism, or else

is violating the epistemic principle that we should apportion our credences to correspond to the objective chances (cf. Lewis 1986). But this reason doesn't carry over to the normative case, where the truth of the matter was not settled by a chance process. It's only by employing our normative judgment that we can draw any conclusions about the likelihoods of various normative claims (Schafer 2010). That's the most important point. But the following consideration is also of interest: It's not entirely clear that the irrationality of the physical lottery player consists in *internal incoherence*; we may instead take them to be *substantively* irrational, as a matter of mind-independent normative fact, though of course Street cannot accommodate such a "realist" thought.

So I think Street's argument strictly fails: There's nothing incoherent about assigning higher odds (a priori) to anti-pain normative systems than to pro-pain views. So the moral realist needn't conclude that she is "probably" among the misguided when she holds to her anti-pain view. It's true that our views are structurally analogous to others (like the pro-pain view) that we hold to be mistaken. But there's nothing wrong with the *structure* of those other views — they're just wrong on the substance. So we shouldn't be particularly troubled to share the structure of the wrong-headed view, just with different substance. That's how the right view would *have* to look.

Suppose you grant that the realist *needn't* (on pain of incoherence) take herself to be "probably" among the misguided. Still, you may think, perhaps it's still the case that she really *should*, as a matter of substantive epistemic

rationality, conclude this. But now you're positing a mind-independent normative truth. And if the argument against normative realism depends upon the truth of realism, then that can't possibly be a sound argument!

2   A Positive View

2.1   Explication

We've seen that there's no sound argument against realism to be found here. But there is at least a *challenge* for the realist, to say more about how moral knowledge is possible on her view. I can't give a full answer here, but here's a rough sketch of a view:[6] There's a fact of the matter as to which psychologies qualify as *fitting* or 'substantively rational'. The fitting psychology might, among other things, endorse modus ponens as a valid rule of inference, accept inductive over counter-inductive norms, and take the badness of pain as a provisional moral datum. More generally: the fitting psychology is one that reflects or "fits with" the objective normative facts, whatever they may be. So, if the truth of p would in fact be a reason to believe that q, then the fitting psychology is one that responds to the appearance of p with an inclination towards believing q. If some outcome o is intrinsically desirable, then the fitting agent is one who intrinsically desires o. Inferences drawn, and conclusions reached, by agents with fitting psychologies are, on this view, thereby justified — and, if true, eligible to qualify as knowledge.

---

[6] I don't mean to suggest that the view I go on to sketch is the *only* option for the realist.

That seems a coherent view, and one that allows for (fitting) agents to know the mind-independent moral facts. One might think of it as akin to a kind of reliabilism (the view that true beliefs formed via reliable methods thereby constitute knowledge). Simple reliabilism is clearly immune to Street-style skeptical arguments: The mere fact that there are *other*, less reliable, processes does nothing at all to undermine the process that I actually use. Even if I have no independent way of verifying that the process I use is one of the few reliable ones, all that matters is that it *is* reliable as a matter of actual fact. And this highlights the disanalogy with the physical lottery: Presumably there's no way to be a reliable predictor of lottery results, the way that an anti-pain moralist might be a reliable predictor of moral truths. Similar observations may be made when using the fittingness view in place of reliabilism.

One reason to prefer the fittingness view to traditional reliabilism is that the latter implausibly holds the justification of our beliefs hostage to purely external circumstances. The sensory "perceptions" of a Brain in a Vat, for example, are not very reliable, but surely my BIV-duplicate is no less rationally justified in his beliefs than I am in mine (Cohen 1984). The fittingness view, by contrast, is "internalist" in the following important sense: Epistemic justification supervenes on one's mental states, so that any two intrinsic duplicates are alike with respect to the justificatory status of their beliefs. We might characterize the fittingness view as *a priori expected reliabilism*, whereby what matters is not whether my actual belief-forming mechanisms

are reliable in the actual world, but rather, whether my actual belief-forming mechanisms are *expectably* reliable, given the objective a priori probability distribution over possible worlds.[7] (This probability distribution, being a priori, does not differ depending on which world is actual. This is how the view manages to be "internalist" in the above sense.)

The fittingness view can also comfortably grant that *mere* reliability is not enough.[8] For example, an agent who merely found themselves automatically parroting the claim that slavery is wrong, without any understanding of the reasons why, might well strike us as a poor candidate for moral knowledge. But the fitting agent is not so vacuous. Rather, she possesses a complex and interweaving 'web' of moral beliefs, accurately reflecting the explanatory relations that hold between the actual moral facts. She thus understands *why* slavery is wrong, such that she is reliably accurate not just on this precise question but also various permutations of the case. She knows what the morally relevant considerations are, and exhibits sensitivity to them in two crucial respects: (1) she recognizes when similar underlying moral considerations arise in superficially different cases, and (2) she recognizes when morally relevant differences should prompt different verdicts in superficially similar cases. Fill in further details of the fitting or 'substantively rational'

---

[7] The idea of an a priori objective probability distribution is controversial, but I think it can do important explanatory work. The basic idea is that some possible worlds were, as a matter of brute fact, *more likely* to be actualized than others. For example, we may think that simpler and more uniform worlds are more 'eligible' or likely to exist, and this is why it is fitting to reason according to inductive rather than counter-inductive norms.

[8] Thanks to William Fitzpatrick for pressing me on this point.

psychological profile as you please. What matters for my purposes is, firstly, that such a psychological profile could emerge in an agent through purely natural causes, and secondly, that the true beliefs of such a broadly reliable — indeed, *wise* — moral agent plausibly constitute moral knowledge.

To wrap up: It may be epistemically undermining if *this psychology* I'm using was lucky to get things right. That casts doubt on the rationality of my psychological processes. But there's no such problem if I'm merely lucky to have *acquired* a rational psychology in the first place. That doesn't cast doubt on the rationality of the psychology I have; it merely suggests that I could easily have ended up with some other, less rational, psychology instead. To which the appropriate response is simply, "Thank goodness I didn't!"

Non-skeptical normative realists are merely committed to the latter, unproblematic kind of epistemic luck. Given the fact of coherent moral diversity, and the causal irrelevance of the moral facts, there's no guarantee that agents will end up with substantively rational or normatively fitting psychologies. And there's no "neutral", purely procedural way to test for substantive normative correctness. All that we can do is pursue wide reflective equilibrium, resolving the inconsistencies in our thoughts in whatever way strikes us as overall most plausible, and hope for the best.[9] When we consider all the possible ways of being crazy without realizing it, we may have a feeling of "There but for the grace of God go I." In this sense, we must consider our-

---

[9] That's not to say that reflective equilibrium is sufficient. What we really need, to be in a good epistemic position, is to pursue reflective equilibrium *from an objectively adequate starting point.* Cf. Kelly and McGrath (2010).

selves "lucky" to be rational: We can just as well imagine circumstances causing us to have substantively crazy normative commitments, and it was not in our control that we happened to be born into the good circumstances rather than the bad. But, crucially, none of this implies that we are *now* unreliable, or that *given our actual psychology* the odds of our answering moral questions correctly is low. Being in some sense "lucky" to possess a reliable, rational, and competent psychology does not, of course, make one's actual psychology any less reliable, rational, or competent. It most certainly does not imply that the correctness of one's conclusions is mere luck or accident, as opposed to being the expected result of one's rational competence. But that is the conclusion the skeptic needs to establish in order to undermine normative realism. Since they cannot establish this, the skeptical challenge is thus defanged.

2.2 Defense

While I haven't space here to offer a full-fledged defense of the above sketched view, let me offer a few quick comments to assuage the concerns of some who might initially regard it as excessively "dogmatic". So far I've shown that the view is internally coherent and defensible, even in the face of Street's arguments. I now want to suggest that many should find the view positively appealing, as the best way to render coherent our various common-sense commitments. Here are three considerations in support of this conclusion:

First, there is the Moorean point that we may reasonably be more con-

fident in our first-order moral views than we are of any skeptical principle to the contrary, including Street's principle (from §1.3, that we should regard our moral faculties as unreliable unless we can give a constitutive, as opposed to substantive, explanation of their reliability). Street hopes that her constructivist anti-realism can accommodate enough of the moral data to provide a viable alternative to moral realism. I think it's far from clear that we should have greater credence in Street's principle than in the realist's datum that even pro-pain agents shouldn't engage in gratuitous torture. Regardless, as I'll now argue, we can rule out Street's constructivism on the stronger grounds that it is self-defeating.

Constructivism is the view that eligible normative judgments — those that "withstand scrutiny" from the agent's normative standpoint — thereby constitute normative truths (Street ms). Call $S_{ProPain}$ the situation of making an eligible judgment that one ought to engage in gratuitous torture. Constructivism implies that any agent in $S_{ProPain}$ thereby really ought to engage in gratuitous torture. Yet many realists hold the contrary belief that they shouldn't engage in gratutious torture *even if* they found themselves in situation $S_{ProPain}$. If this normative belief of ours also withstands scrutiny, then constructivism will imply that it is true, thereby contradicting itself. Constructivists are thus committed to showing that moral realism is incoherent, or cannot withstand scrutiny from the normative standpoint of any possible agent.[10] But we've seen that moral realism is *not* incoherent (however

---

[10] Street (ms, 36) explicitly accepts this point, though William Fitzpatrick has pointed

substantively implausible some may find it), and so Street's constructivist alternative fails. This forces us back to the initial choice between flat-footed realism and skepticism, reinforcing the Moorean case for the former.

Second, the epistemologically awkward feature of moral realism — that any explanation of our reliability must presuppose substantive claims within the domain — is not unique to the moral domain. It seems equally true of our inductive norms, for example. We can imagine an internally-coherent "counter-inductivist" who takes the sun's rising in the past as evidence that it will *fail* to rise in the future. There is not enough common ground to productively argue with a person who has such fundamentally different epistemic norms. Still, outside of the philosophy seminar room, few of us take the problem of induction at all seriously. The possibility of alternative inductive practices does not seem to undermine our confidence in our own inductive practices. If the possibility of fundamental epistemic diversity does not undermine our epistemic norms, it's unclear why the moral case should be so different.

Third, a general lesson from reflecting on radical skepticism is that we sometimes need to make our peace with having reached justificatory bedrock.

---

out to me that a constructivist might avoid this challenge by restricting "eligibility" to *judgments about one's own actual current circumstances.* But this would leave ungrounded the various normative truths about non-actual, non-current circumstances. If the constructivist offered their theory as a mind-independent account of what's *objectively* true of those cases, then their theory would fall victim to all the same arguments that Street lobbied against moral realism. As Street acknowledges, her arguments can only support a more thorough-going constructivism: one that holds itself to be the unique metaethical truth that would be constructed from any possible standpoint.

We feel the pull of the Socratic ideal to "question everything", and to defend our views against all comers with non-question-begging reasons. But this simply cannot be done (cf. Railton 2004), so we should not grant too much weight to our persisting sense of unease over this.

3 Actual and Possible Disagreement

So far I've argued that Parfit is too concessive to causal-origins-based skepticism. In this section, I will argue that he concedes too much to disagreement-based skepticism. In particular, I'll argue that Parfit's response to merely possible disagreement should also extend to relevant cases of actual disagreement. To support this argument, I'll offer an analysis of when it makes an epistemic difference for a disagreement to be actual, and show that the relevant moral cases do not possess this feature.

3.1 Parfit on Disagreement

Parfit characterizes the *argument from disagreement* as follows:

> If we had strong reasons to believe that, even in [procedurally] ideal conditions, we and others **would** have deeply conflicting normative beliefs, it would be hard to defend the view that we have the intuitive ability to recognize some normative truths. We would have to believe that, when we disagree with others, it is only we who can recognize such truths. But if many other people,

even in ideal conditions, could not recognize such truths, we could

not rationally believe that we have this ability. How could *we* be

so special? (*OWM*, II, 546, bold added)

Parfit is clearly concerned about this argument, and immediately concedes that "Intuitionists must defend the claim that, in ideal conditions, we and others *would not* have such deeply conflicting beliefs." (546, emphasis added) He spends the rest of the chapter defending this convergence claim, following in the vein of the first volume of *On What Matters* (which argued for the convergence of Consequentialism, Contractualism, and Kantianism).

As noted previously, Parfit does grant the fact of *coherent moral diversity*, that people *could* have radically divergent moral beliefs even in procedurally ideal conditions. But such *merely possible* disagreement does not bother him nearly as much. "If we claim that we have some ability," he writes, "it is no objection that we might have lacked this ability." (545) This response is very much in the spirit of the epistemology I sketched in the previous section: What matters is just that we in fact possess a rationally fitting psychology, and are thereby capable of recognizing self-evident moral truths as such. So long as our moral belief-forming processes are indeed a "fairly reliable way of reaching the truth" (545), it is of no matter that we — or others — could have possessed some *other*, less reliable, psychology. (It can be a little unnerving when we reflect on the fact that there's no non-question-begging way to *show* that our psychology is as reliable as we hope it is, but so it goes.)

But now note that this reliabilist response doesn't depend upon the non-actuality of the "other, less reliable psychologies" in question. They could be actually realised — whether in our neighbours or distant aliens — and it would still make no difference to the reliability of our moral beliefs (assuming that we don't defer to those misguided others). If it's reasonable to consider ourselves morally reliable or "special" compared to a hypothetical pro-pain agent, what difference should it make if the imagined agent turns out to actually exist? Conversely: If it's epistemically undermining to be faced with an internally coherent alternative to your present views, why should it matter whether the advocate of this alternative view really exists, or is merely a figment of your imagination playing devil's advocate?

One possibility is that Parfit is thinking of our moral belief-forming process in very coarse-grained terms, e.g., as "[b]elieving what seems self-evident, after [careful] reflection" (545), a process that might be shared by pro-pain and other thoroughly wrongheaded agents.[11] So whether this shared process qualifies as statistically "reliable" or not depends on what proportion of actually existing agents have the right intuitions to begin with. If there turns out to be great diversity in what moral beliefs actually result from using this process, then it doesn't reliably yield *any* particular results, let alone reliably *correct* ones.

---

[11] Ashley Atkins and Sarah McGrath have suggested, in conversation, that a more plausible variant of this view might be restricted in scope to some more limited moral community, e.g. just other humans, leaving distant aliens aside. But my below objection still applies to this more moderate view.

However, as suggested in §2.1, the rational status of our beliefs should not depend in this way on how things are outside of our heads. This untoward consequence is avoided by more fine-grained approaches, such as the fittingness view, which see our starting beliefs and intuitions as an essential part of our overall belief-forming process. This means we should separately assess the reliability of pro-pain and anti-pain moralists, rather than indiscriminately lumping them all together into a single class of epistemically-equivalent agents. Given this more fine-grained categorization, we avoid the absurd result that the reliability of *our* moral belief-forming process depends on the absence of creatures with radically *different* intuitive starting points.

3.2   When Actual Disagreement Matters

There are clearly some cases in which actual disagreement matters. If we learn that my thermometer actually disagrees with yours about the temperature, that's epistemically significant in a way that merely imagining a divergent thermometer reading is not. But I think the sort of "ideal disagreement" that metaethicists are interested in does not share this feature. So let me say a bit more about these two kinds of disagreement, and why the modal status of a disagreement (as actual or merely possible) makes no difference in cases of the second kind.

In cases of what we might call 'non-ideal' disagreement, there's a presumption that the disagreement is rationally resolvable through the identification of some fallacy or procedural mis-step in the reasoning of either ourselves

or our interlocutor. The disagreement is 'non-ideal' in the sense that we're only disagreeing because one of us made a blunder somewhere. We are sufficiently similar in our fundamental epistemic standards and methods that we can generally treat the other's output as a sign of what we (when not malfunctioning) would output. The epistemic significance of the disagreement is thus that the conflicting judgment of a previously-reliable source is some evidence that we have made a blunder *by our own lights*, though we may not yet have seen it. Now, obviously, merely *imagining* a blunder-detector going off in our vicinity is no more evidence of an actual blunder on our part than is an imagined fire siren evidence of an actual fire.

The case of 'ideal' or irresolvable disagreement is rather different. In this case, both agents are (we may stipulate) logically omniscient and hence fully confident that they have not made any procedural blunder in their reasoning. The other's disagreement casts no doubt on this, because the disagreement is instead traceable to a much more fundamental divergence in starting assumptions.

Here the epistemic significance of the disagreement is more indirect. It's significant just in that it brings to our attention a fact that we might otherwise have neglected: *there's this internally coherent alternative worldview against which we can muster no non-question-begging argument*. But of course we might just as well be gripped and troubled by this fact even if there never actually existed any advocate for the view in question. The challenge is just: why accept our worldview rather than some other? Any answer

we try to give will naturally draw on the assumptions of our own worldview, and hence prove dialectically unsatisfying. But given the inevitability of this, perhaps it shouldn't bother us too much.

Conclusion

We've seen that Parfit addresses skeptical arguments, whether causal-origins-based or disagreement-based, by conceding the first step to the skeptic and then trying to draw a line further down the track. But this proved to be a slippery undertaking. In the causal origins case, Parfit draws a line between evolutionary influences on belief and other causal influences, claiming that only the former are sure to be epistemically undermining. In the disagreement case, he distinguishes actual and possible disagreement, again claiming that only the former should concern us. In both cases, I have argued, these distinctions lack the epistemic significance Parfit attributes to them, and leave the realist needlessly vulnerable to empirical contingencies. If we are to defend moral realism against these skeptical challenges, we must instead address them at their first step. By appealing to a new kind of reliabilist moral epistemology, I hope to have shown how we can meet this deeper challenge.

References

Cohen, Stewart. 1984. "Justification and Truth." *Philosophical Studies* 46:279–95.

Copp, David. 2008. "Darwinian Skepticism About Moral Realism." *Philosophical Issues* 18:186–206.

Enoch, David. 2010. "The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It." *Philosophical Studies* 148:413–438.

Kelly, Thomas and McGrath, Sarah. 2010. "Is Reflective Equilibrium Enough?" *Philosophical Perspectives* 24:325–359.

Lewis, David. 1986. "A Subjectivists Guide to Objective Chance." In *Philosophical Papers*, volume 2. Oxford: Oxford University Press.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

—. 2011. *On What Matters*, volume 2. Oxford: Oxford University Press.

Railton, Peter. 2004. "How to Engage Reason: The Problem of Regress." In R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds.), *Reason and Value: themes from the moral philosophy of Joseph Raz*. Oxford: Clarendon.

Schafer, Karl. 2010. "Evolution and Normative Scepticism." *Australasian Journal of Philosophy* 88:471–488.

Street, Sharon. ms. "Objectivity and Truth: You'd Better Rethink It."