1. Introduction

Moral disagreement is pervasive among both philosophers and non-philosophers. Perhaps it’s no surprise that non-philosophers disagree about ethics, but widespread, intractable moral disagreement among philosophers is more striking (and more troubling). Aren’t moral philosophers supposed to be the experts—maybe not with respect to being moral but at least with respect to knowing the moral truths? What do ethicists do at their jobs all day if they’re not discovering, or making progress toward discovering, the moral truths? And why, after several thousand years of ethical inquiry, have they not been able to make any progress toward consensus? It seems philosophers are no closer to reaching agreement about ethics today than they were when Socrates walked the streets of Athens. Indeed, dissensus seems to have grown since then.

These reflections can cause one to have doubts about morality. One might doubt, first, that many of our moral beliefs are epistemically justified: If the people who devote their lives to ethical inquiry can’t come to agreement about ethics—and show no signs of ever doing so—then how can we (or they) be confident that our (or their) moral beliefs are true? One might even doubt that there are any moral truths at all. After all, one possible explanation for the pervasiveness and intractability of moral disagreement among moral philosophers is that there are no moral truths about which to agree. Perhaps disputes about how we ought to live are no better than those disputes in ancient times about whether the moon is male or female---there’s just no fact of the matter. These kinds of doubts find their expression in arguments from disagreement in metaethics.
Arguments from disagreement are typically aimed at *non-skeptical moral realism*. As I will understand this view, it comprises four claims:

**Cognitivism:** moral judgments express beliefs that aim to represent moral reality,

**Non-nihilism:** some of these positive moral judgments are true,

**Stance-independence:** the moral truths do not constitutively depend on the attitudes of any actual or hypothetical agent(s), and

**Non-skepticism:** we are justified in believing some moral propositions.

While arguments from disagreement might be employed to target other views, such as constructivism or expressivism, I will only be interested in considering the challenge these arguments pose for non-skeptical realism.

Arguments from disagreement begin by calling attention to (or supposing) widespread, fundamental moral disagreement among a certain group of people (e.g., the folk, moral philosophers, idealized agents). Then, some skeptical or anti-realist-friendly conclusion is drawn. These conclusions come in three varieties: epistemological (e.g., we don’t know any moral propositions), metaphysical (e.g., there are no moral truths), or semantic (e.g., moral judgments express non-cognitive attitudes). This paper focuses only on the first two kinds of arguments. Moreover, it focuses on arguments that take as their starting point a particularly troubling kind of disagreement: disagreement among excellent moral inquirers.

For many years now, moral realists have argued that we shouldn’t be terribly concerned about moral disagreement in society, or among “the folk.” Moral disagreements in society (or across societies), realists often say, are the product either of ignorance about the non-moral facts, genuine moral indeterminacy, failures of imagination and sympathy, the distorting influences of self-interest, 

---

1 I’ll call it “non-skeptical realism” or just “realism” from now on.
2 A positive moral judgment is one whose truth entails either that it would be good (or bad) for an agent to do something, that an agent morally ought (or ought not) to do something, that an agent has moral reason to do (or not to do) something, or something similarly unacceptable to moral error theorists.
3 Much of what I say in defense of non-skeptical realism can, however, be adopted by expressivists and constructivists.
or some other kind of cognitive bias or shortcoming. The standard realist line, then, is that moral disagreement in society persists because most people are bad (or less-than-ideal) moral inquirers.\(^4\) If this is correct, then the more troubling kind of disagreement for non-skeptical realists is the kind we see among moral philosophers or the kind we might expect to see between perfectly informed, perfectly rational agents. After all, moral philosophers have more time than almost anyone to investigate the morally relevant empirical facts, try to rid themselves of bias, try to imagine the plight of others, and so on. And ideally informed, ideally rational agents wouldn’t be subject to any of the cognitive shortcomings to which realists often appeal to explain disagreement in society. So, if disagreement in these ideal (or nearly ideal) conditions persists, then realists cannot respond to concerns about moral disagreement in the standard way.

When realists do respond to arguments from disagreement, those responses tend to be piecemeal.\(^5\) They take each argument from disagreement one at a time and try to respond accordingly. In this paper, however, I argue that arguments from disagreement share a structure that makes them vulnerable to a single, powerful objection: they self-undermine. For each formulation of the argument from disagreement, at least one of its premises casts doubt either on itself or on one of the other premises employed on the way to the argument’s skeptical or anti-realist conclusion.

Upon reflection, this should not be surprising. Arguments from disagreement are *philosophical* arguments. They therefore employ philosophical—specifically metaphysical and epistemological—premises to support their conclusions. But deep, widespread, intractable disagreement about philosophy (especially metaphysics and epistemology) is pervasive, both among philosophers and non-philosophers.\(^6\) Thus, any non-trivial metaphysical or epistemological premise an argument from

\(^4\) This popular line is taken by Brink (1987: Ch. 7), Enoch (2011: Ch. 8), and Shafer-Landau (2003: Ch. 9).
\(^5\) Again, this is the strategy employed by Brink, Enoch, and Shafer-Landau in their defenses of realism. I think that their responses are largely correct. Thus, everything I say in this paper is meant to be compatible with those defenses.
\(^6\) Shafer-Landau makes this point forcefully in (2003: 220). He then argues that arguments from disagreement for anti-realism in ethics overgeneralize, committing their proponents (implausibly) to anti-realism about philosophy. I suspect that some arguments from disagreement do fall to this objection, but not the ones I consider in this paper.
disagreement may employ is likely to be the subject of deep, widespread, intractable disagreement. And if the existence of such disagreement about a moral proposition casts doubt on that proposition or on our moral beliefs more generally, as arguments from disagreement purport to show, then, by those arguments’ own standards, such disagreements cast doubt on those arguments’ own controversial premises. By those arguments’ own standards, then, they fail.7

I don’t have space to consider every (metaphysical and epistemological) argument from disagreement. I will therefore show how the self-undermining objection applies to some of the most recent and most powerful arguments, which I take to be representative of many others. I’ll consider two kinds of epistemological arguments and two kinds of metaphysical arguments. In each case, I argue, these arguments employ premises that are highly controversial, and in precisely the same ways that the moral propositions they purport to undermine are controversial. The result is that these arguments undermine themselves. Once I’ve shown how the self-undermining objection applies to the arguments I consider in this paper, it should be clear how the objection will apply to other arguments from disagreement not discussed here.

If the argument presented in this paper is sound, it provides realists a single, unified strategy for responding to arguments from disagreement. Moreover, it provides a challenge for any future arguments from disagreement: any future argument must rely on premises that are not themselves controversial in the way the moral propositions they seek to undermine are. But, again, given the pervasiveness of disagreement among both philosophers and non-philosophers about metaphysics and epistemology, this will be an incredibly difficult task. Thus, I conclude, realists should not be

---

7 Enoch (2011: 215-16) raises a self-defeat worry, too. He argues that arguments from disagreement seek to establish a metaethical position. But, given the level of disagreement in metaethics, arguments attempting to show that there is no fact of the matter about ethics would also show that there’s no fact of the matter about metaethics. And one cannot coherently defend an argument for a metaethical position if that argument also supports the conclusion that there is no fact of the matter about metaethics. Again, there are likely some arguments that fall to this objection, but not the ones I consider here.
concerned about existing arguments from disagreement, and they should rest assured that any future arguments will have an extraordinarily high bar to meet.

2. Arguments from Conciliationism

The first kind of argument I'll consider draws on the peer disagreement literature. The central question in that literature is this: How should we revise our doxastic attitude(s) toward a proposition $p$ when we discover that we disagree about $p$ with an epistemic peer—i.e., someone we judge to be just as intelligent, well-informed, open-minded, and otherwise epistemically virtuous with respect to $p$ as we are? One popular view is

**Conciliationism:** If an agent $A$ learns that an apparent epistemic peer, $B$, disagrees with her about $p$, then $A$ is rationally required either to suspend judgment about $p$, adjust her credence in $p$ significantly in the direction of $B$'s, or both.  

If Conciliationism is the correct view about how to respond to peer disagreement, then it has clear implications for our moral views. Given that many of our moral beliefs are the subject of disagreement among excellent moral philosophers (some of them our epistemic peers or superiors), Conciliationism implies that we are rationally required to suspend judgment about a large proportion of our moral views. There are, after all, excellent philosophers on virtually every side of every moral controversy. If so, then here is a very incomplete list of the moral issues about which Conciliationism would require us to suspend judgment: consequentialism vs. non-consequentialism; whether the doctrine of double effect is true; whether the killing/letting die distinction is morally significant; the moral (im)permissibility of eating meat, abortion, capital punishment, genetic enhancement, prostitution,

---

8 For simplicity, I will mostly speak in terms of full (dis)belief, rather than credences.
and many, many other debates in moral and political philosophy. The complete list would be enormous.10

Armed with Conciliationism, one can advance

*The Argument from Conciliationism*11

1. If Conciliationism is true, then we’re rationally required to suspend judgment about any moral claim that is the subject of disagreement among our epistemic peers.
2. Conciliationism is true.
3. So, we’re rationally required to suspend judgment about any moral claim that is the subject of disagreement among our epistemic peers.
4. A very large proportion of our moral beliefs are the subject of disagreement among our epistemic peers.
5. Therefore, we are rationally required to suspend judgment about a very large proportion of our moral beliefs.

If this argument succeeds, it would not entail that an extreme version of moral skepticism, according to which peer disagreement requires us to suspend judgment about *all* our moral beliefs, is correct. Moral philosophers don’t disagree about everything. There is very little disagreement about, for instance, whether pain is bad, whether a person’s hair color alone counts as a reason to harm them, whether a crying infant deserves to be punished for waking you up from your nap, and so on. These beliefs would not be threatened by Conciliationism. Nevertheless, the truth of Conciliationism would entail that we are required to suspend judgment about most of our interesting moral and political views. So the argument entails a kind of skepticism (or agnosticism) that is troubling enough to warrant a response.

Some have tried to resist *The Argument from Conciliationism* by attacking (1)—the claim that Conciliationism entails that we ought to suspend judgment about a large proportion of our moral beliefs.

---

10 You might, like me, think it’s plausible that we ought to suspend judgment about some of these issues. But I think we ought to suspend judgment because the first-order evidence is inconclusive, not because philosophers disagree about these issues.

11 A version of this argument is advanced by Matheson (2016), McGrath (2007), Rowland (2017), and Sinnott-Armstrong (2007).
views. Others have attacked (4)—the claim that we have many epistemic peers who disagree with us about moral issues—by suggesting that, as a matter of fact, we have very few epistemic peers about moral matters. I don’t find either strategy compelling, but I will not argue here that they fail. Instead, I’ll argue that no one, not even conciliationists, should accept this argument because (2)—the claim that Conciliationism is true—is self-undermining. If I am correct, then our justification for resisting the Argument from Conciliationism doesn’t depend on the cases others have made for rejecting (1) and (4).

Conciliationism is a popular view in the epistemology of peer disagreement, but it hasn’t won universal acceptance. Many epistemologists hold a version of a competing view, known widely as the Steadfast View (see, e.g., Bergmann (2009), Decker (2014), Enoch (2010), Kelly (2005), Titelbaum (2014), van Inwagen (2010), Weatherson (2013), Wedgewood (2010), Weintraub (2013)). According to steadfasters, disagreement about p between epistemic peers is rarely sufficient, all by itself, to rationally require either peer to suspend judgment about p. Thus, Conciliationism is itself the subject of disagreement among excellent philosophers. This fact poses an obvious problem for conciliationists. If, as virtually all conciliationists admit, conciliationists have many epistemic peers who disagree with them about the truth of Conciliationism, then it seems to be irrational, by conciliationists’ lights, for conciliationists to believe (or be confident in) their own view. Call this the self-undermining problem (SUP).

Conciliationists are, of course, aware of this problem. Some of them accept without reservation the obvious implication that we cannot rationally believe Conciliationism. For instance, David Christensen, a prominent conciliationist, writes

---

12 See, e.g., Vavova (2014).
13 See, e.g., King (2012).
14 As with Conciliationism, the Steadfast View is best described as a family of views.
15 I’m not the first to point out that arguments from Conciliationism face the SUP. Decker and Groll (2013) and Horn (2016) do this, too. But they don’t engage conciliationist attempts to justify confidence in Conciliationism even in the face of peer disagreement, as I do here.
It seems to me those of us who find ourselves strongly drawn toward Conciliationism in these contentious times should not be confident that Conciliationism is correct. (Of course, we may still work hard in producing and disseminating arguments for the view, hoping to hasten thereby the day when epistemic conditions will brighten, consensus will blossom, and all will rationally and whole-heartedly embrace Conciliationism.) (2009: 763)

And later he writes

[Recognition of the SUP] puts the advocate of CV [conciliatory views] in a situation that's puzzling in a number of ways. For one thing, it would seem that, in the present epistemological climate, at least, CV has the property that one cannot rationally believe it (at least very strongly), even if it's true. (2013: 78).

Christensen later notes, correctly, that the fact that conciliationists cannot rationally believe their view, given the current controversy about it, does not entail that Conciliationism is false. Nevertheless, he does concede that Conciliationism cannot now be rationally believed, and this is what concerns me now. I will not (and need not) argue that Conciliationism is false to show that we should not accept the Argument from Conciliationism. It's sufficient, for my purposes, to show that the Argument from Conciliationism employs a premise we cannot rationally believe. After all, no one should accept the conclusion of an argument on the basis of that argument if, by one's own lights, one cannot rationally believe one of the premises.

While Christensen doubts that we can rationally believe Conciliationism, other conciliationists are more optimistic. They argue that we are rationally permitted to believe Conciliationism, even in the face of peer disagreement about it. If they’re correct, then conciliationists can escape the force of the SUP and Conciliationism wouldn’t self-undermine in the way I’ve described. But, as I’ll now argue, these defenses fail for roughly the same reason. Each defense depends for its success on philosophical claims that are themselves the subject of disagreement among conciliationists’ epistemic peers. It is therefore irrational, by conciliationists’ own lights, to believe the claims they employ in their defenses.
It is thus irrational to believe that these defenses succeed and, given the SUP, irrational to continue believing Conciliationism in the face of peer disagreement about it.16

Begin by considering Tomas Bogardus’s (2009) defense of Conciliationism against the SUP. He argues that the truth of Conciliationism is obvious and can be known by direct acquaintance on the basis of rational intuition. Thus, any apparent epistemic peer who denies Conciliationism can be demoted from the status of epistemic peer on the grounds that they deny an obviously true proposition known by direct acquaintance. Bogardus writes,

[T]he antecedent of the Equal-Weight View [that some agent is an apparent epistemic peer] might not be satisfied in cases involving knowledge from that unmediated access to the truth of propositions sometimes afforded by rational intuition. And it’s plausible that the Equal-Weight View is itself a deliverance of rational intuition . . . With further reflection, I think, one can come to just see the truth of the View – not only does it seem obvious, but upon further reflection it just is obvious (2009: 333).

Clearly, this defense succeeds only if the following claim is both true and can be rationally believed.

**Obvious**: The Equal Weight View—one version of Conciliationism—is obviously true and is known by direct acquaintance on the basis of rational intuition.

On Bogardus’s view, Obvious serves as the justification for demoting any steadfast from the status of epistemic peer. But one can rationally demote a peer on the basis of a belief in Obvious only if one can rationally believe Obvious. Thus, the success of Bogardus’s defense depends on the claim that Obvious is both true and can be rationally believed. Unfortunately for conciliationists, however, Obvious cannot, by conciliationists’ standards, be rationally believed since it is itself highly controversial. Not only do all steadfasters deny Obvious, but so also do many conciliationists.

One might suggest that Obvious, too, is obvious or known on the basis of rational intuition. If this were true and we could rationally believe it, then conciliationists could demote any apparent epistemic peer who disagrees with them about Obvious on the grounds that they deny a proposition

---

16 I argue for this conclusion in much more detail in (Sampson, ms.). What follows is a rough sketch of that argument.
that is obviously true. But the claim that Obvious is obvious or known on the basis of rational intuition is even more controversial than Obvious itself. Again, every steadfaster will deny this and so also will many conciliationists. Conciliationism thus entails that it’s not rational to believe that Obvious is obvious or known on the basis of rational intuition. So one cannot rationally demote an apparent epistemic peer on the basis of a belief that Obvious is obvious. One might attempt to make the same move yet again—i.e., insist that it’s obvious that Obvious is obvious. But I take it that no conciliationist would want to pursue this desperate strategy.

The same difficulty afflicts Adam Elga’s (2010) famous defense of Conciliationism against the SUP. Elga employs a self-exempting strategy to defend Conciliationism: He argues that we rationally ought to conciliate in the face of peer disagreement, except when the disputed proposition is Conciliationism itself. The resulting view is

**Elga’s Self-Exempting Conciliationism (ESEC):**

(1) Conciliationism: If \(A\) learns that an apparent epistemic peer, \(B\), disagrees with her about \(p\), then \(A\) is rationally required to adjust her credence in \(p\) significantly in the direction of \(B\)'s,

(2) unless \(p\) is Conciliationism. If \(p\) is Conciliationism, then \(A\) ought to remain steadfast in her confidence—indeed have a credence of 1—in \(p\).

Obviously, Elga needs to explain how ESEC’s second clause isn’t objectionably *ad hoc*. Elga’s answer is that Conciliationism—ESEC’s first clause—is an inductive method. And *all* correct inductive methods must be dogmatic—i.e., recommend credence 1—about their own correctness. If an inductive method were not dogmatic about its own correctness, Elga argues, then there would be possible circumstances in which the method calls for its own rejection. But that would render the method incoherent. It would be as if the method were to say, “Take the following advice: do not take this advice.” That’s incoherent. Elga thinks that any method that could possibly give such incoherent advice is a false (or bad, or incorrect) method. So, on Elga’s view, for *any* correct inductive method
M, M will never advise an agent to stop believing or taking advice from M. That includes whatever method ought to govern our response to peer disagreement. Thus, Elga argues, his self-exempting Conciliationism is not objectionably ad hoc. It treats Conciliationism like any other inductive method. If this defense is correct, and we can rationally believe it, then he’s successfully rebutted the SUP.

Unfortunately for Elga, whether his defense succeeds is the subject of disagreement, not only among steadfasters (who obviously believe that his defense fails), but among his fellow conciliationists, too. For example, David Christensen has criticized Elga’s self-exempting Conciliationism at length (2013: 88-89). Christensen’s criticism, in short, is that, to endorse Elga’s defense by adopting credence \(1\) in Conciliationism, one would be committed to believing either that one is infallible when it comes to reasoning about peer disagreement or that one is just incredibly lucky to have hit upon the correct view. But neither suggestion is plausible. So, according to Christensen, we cannot rationally endorse ESEC. If Christensen is correct, then Elga’s response cannot insulate Conciliationism from the force of the SUP.

The point of the preceding paragraph is not to argue that Christensen’s criticism is correct (though I’m inclined to think that it is). Rather, the point is that the question of whether Elga’s response succeeds in defending Conciliationism from the SUP, is itself the subject of disagreement among excellent philosophers—not just among steadfasters, but among conciliationists, too. By Elga’s own lights, then, it is irrational to believe that his defense succeeds. It would therefore be irrational to believe Conciliationism in the face of the SUP on the basis of Elga’s defense. So, unless some other defense of Conciliationism against the SUP comes to the rescue, conciliationists cannot rationally believe their view. Perhaps such a defense will come, but it doesn’t exist now.\(^{17}\)

\(^{17}\) I argue elsewhere (Sampson, ms), however, that it is highly unlikely that any such defense will emerge in our lifetimes, given the extent of disagreement in epistemology, especially about the standards of rational belief.
The upshot, then, is that no one should believe the conclusion of the Argument from Conciliationism on the basis of that argument, since, given the actual distribution of opinion about Conciliationism, its second premise is self-undermining. No one can rationally believe it. And if no one can rationally believe that premise, then no one can rationally accept the conclusion of the Argument from Conciliationism on the basis of that argument.

3. Arguments from the Unreliability of the Method of Moral Inquiry

Another popular form of epistemological argument against non-skeptical realism exploits moral disagreement to cast doubt on the reliability of the method used in moral inquiry. These arguments begin by calling attention to the troubling fact that moral disagreement among philosophers persists despite philosophers’ use of (what is widely regarded as) the best method for moral inquiry: reflective equilibrium. Reflective equilibrium is, roughly, the process of making intuitive judgments about moral cases and general moral principles, and then bringing those judgments into coherence with one another and with our beliefs about the non-moral facts. While moral philosophers don’t agree about much, it’s remarkable how much they agree that reflective equilibrium is the correct way to proceed with ethical inquiry. This makes it all the more discouraging that this method has failed to lead many, perhaps most, philosophers to the moral truth. After all, moral disagreement is pervasive among philosophers, and it can’t be that all their logically incompatible moral judgments are correct—not on non-skeptical realism, at least. This suggests that even our best method of moral inquiry is unreliable. And it seems clear to many that if one is using a method of inquiry that one knows is unreliable, then one is not epistemically justified in believing the outputs of that method. In the case of moral inquiry, then, it seems that we are not epistemically justified in holding our moral beliefs—the products of reflective equilibrium.
Perhaps the best version of this argument comes from Justin Horn (2016). Horn argues that there are plausible cases of fundamental moral disagreement. A moral disagreement is fundamental if it would persist even if the disputants had all the morally relevant non-moral information, were making no logical mistakes, and were not subject to any other sort of bias. One example of such a disagreement, Horn suggests, is the disagreement about whether certain items of non-sentient nature (e.g., rivers, rock formations, mountain ranges) are non-instrumentally valuable. Some environmentalists suggest that trees, for instance, in addition to being instrumentally valuable (because they give off oxygen and can be made into paper) are also non-instrumentally valuable. Others insist that trees are valuable only insofar as they help contribute to the flourishing of sentient life. It’s not clear that either group is misinformed about some non-moral matter. Nor is it clear that either has made a mistake in reasoning or is subject to some bias. If that’s correct, then the disagreement among philosophers about the non-instrumental value of non-sentient nature is fundamental.

We might wonder what explains fundamental moral disagreement when it occurs. Why do excellent philosophers reach different moral judgments? Horn suggests that such disagreements are plausibly explained by the hypothesis that the two parties have different starting points—namely, conflicting moral intuitions. If so, then the difference in intuitive judgments is systemic such that not even flawless coherentist reasoning could bring the two parties into agreement. Cases of systemic differences in moral intuitions, Horn thinks, raise a serious epistemological challenge for non-skeptical realists. He writes:

If I have a disagreement with an equally rational and (non-morally) informed interlocutor and this disagreement is ultimately traceable to a systemic difference in our intuitions, then at least one of us must have systemically misleading intuitions. Given that we are equally rational and non-morally informed, why should I think that my intuitions happen to reflect the truth while my opponent’s do not? And if there’s reason to think that even our most careful moral judgments fail to reliably track any stance-independent moral reality, is there any remaining reason to believe in such a reality at all? (Horn: 4)

Horn then proceeds to advance what he calls the
Reliability Argument from Disagreement (RAD):

6. If there is widespread fundamental moral disagreement, then many people would have a significant number of false moral beliefs even if they were to flawlessly employ our best method of moral inquiry.
7. If many people would have a significant number of false moral beliefs even if they were to flawlessly apply our best method of moral inquiry, then that method is not reliable for many people.
8. If one is aware that our best method of moral inquiry is not reliable for many people, then one cannot be justified in believing the outputs of such a method unless one has special reason to believe that it is reliable in one’s own case.
9. It is not the case that any of us has special reason to believe that the best method of moral inquiry is reliable in our own case.
10. Therefore, if we become aware of widespread fundamental moral disagreement, then we cannot be justified in believing the outputs of our best method of moral inquiry (Horn: 12).

Three features of the RAD are worth noting. First, the conclusion is conditional. Horn is reluctant to speculate about the extent to which there is actual fundamental moral disagreement. He correctly notes that that’s an empirical question that cannot be settled from the armchair. He is more interested in exploring the implications of such widespread disagreement if it exists. I will, however, grant for the sake of argument that there is such widespread, fundamental disagreement and that we are aware of it. I will therefore proceed as if the conclusion of the RAD is not conditional—that the conclusion is that we cannot be justified in believing the outputs of our best method of moral inquiry.

Second, the RAD purports to undermine the justification we might have for any of our moral beliefs, not just our controversial moral beliefs. This is what distinguishes the RAD from the Argument from Conciliationism. On Horn’s view, reflective equilibrium is our best (perhaps only) method of moral inquiry. Widespread, fundamental moral disagreement reveals that there is something deeply flawed about that method. We therefore ought not to trust that any of the outputs of that method are correct.

Third and finally, the RAD is specifically about moral disagreement. There is, however, no reason to restrict its scope to moral but not epistemic disagreement. Anyone who accepts the RAD
as an argument for moral skepticism should accept it as an argument for epistemic skepticism as well, provided there is widespread, fundamental epistemic disagreement. After all, reflective equilibrium is the best (perhaps only) method we have for epistemological inquiry, too.

We should ask, then: Is there widespread, fundamental epistemic disagreement? It seems that there are disagreements in epistemology, even among the experts, that bear all the same marks of supposed fundamental moral disagreement. Consider the disagreement among epistemologists about the conditions under which a belief is epistemically justified. There have, for decades now, been clearly defined camps, among which are foundationalists, coherentists, reliabilists, infinitists, and others. It doesn’t seem that the members of these camps are lacking some crucial non-normative information, or are subject to some bias, or are making some obvious mistake in reasoning. As with fundamental moral disagreement, it seems that disagreements about epistemic justification are explained by a difference in epistemic intuition between the members of the opposing sides. If so, then the debate about which theory of epistemic justification is correct is an example of fundamental epistemic disagreement.

There are other examples. There are long-lasting debates about, among many other issues, internalism vs. externalism about epistemic justification, whether knowledge is closed under known entailment, whether skepticism can be refuted, whether there is a priori knowledge, and more recently, whether one ought always to conciliate in the face of peer disagreement, and whether a body of evidence permits differing doxastic attitudes toward a proposition for an agent (or between different agents). There’s no more reason, that I can see, to think that these epistemological debates could be settled by providing the disputing parties with more non-normative information, removing biases, or correcting mistakes in reasoning than there is in the case of moral disagreement. If anything, epistemological disagreements seem less likely than moral disagreements to be the product of mistaken non-normative information or bias. After all, the practical stakes seem much lower in these
epistemological debates than in the moral debates. Thus, epistemological disagreements aren’t as easily explained away by the distorting influence of self-interest, or a failure of imagination or empathy, or mistaken empirical beliefs.

If this is correct—if, that is, there is at least as much fundamental disagreement in epistemology as there is in ethics—then the RAD succeeds in undermining the justification we have for our moral beliefs only if it also succeeds in undermining the justification we have for our epistemological beliefs. But premises (8) and (9) of the RAD are epistemological claims. So, if the RAD undermines our moral beliefs, then it undermines our justification for believing (8) and (9), too. But if we shouldn’t believe (8) and (9), then we shouldn’t believe the conclusion of the RAD on the basis of the argument, since it employs premises that, even by the argument’s standards, we’re not justified in believing.

Now we have a clear dilemma for proponents of the RAD. Either the RAD is sound or it isn’t. If the RAD is not sound, then we shouldn’t accept the RAD’s conclusion on the basis of the RAD (since the argument is unsound). If the RAD is sound, then we shouldn’t accept the RAD’s conclusion on the basis of the RAD (since the RAD would undermine our justification for believing premises (8) and (9)). Either way, we shouldn’t accept the RAD’s conclusion on the basis of the RAD.

As with the Argument from Conciliationism, I clearly haven’t shown that the RAD is unsound. I’ve shown only that no one can rationally believe the conclusion of the RAD on the basis of the RAD. Some might find this unsatisfying, but all I can say is that I don’t. After all, neither the RAD nor the Argument from Conciliationism purport to show that our moral beliefs are false, only that we cannot rationally hold them. I’ve shown that precisely the same criticism applies to our beliefs about at least one of the premises in the RAD and the Argument from Conciliationism.18 Thus, by those

---

18 The strength of my critique of the RAD and the Argument from Conciliationism is thus exactly proportional to the (supposed) strength of those arguments’ critique of our moral beliefs. This should not be surprising since my claim is that
arguments’ own standards, we cannot rationally endorse either argument as one that supports moral skepticism (or agnosticism).

4. **IBE Arguments from Intractable Disagreement**

I’ll now turn to arguments that draw some troubling *metaphysical* conclusion about morality. One popular form of metaphysical argument calls attention to disagreement among a certain group of people and suggests that the best explanation for the existence and character of that disagreement is that there are no moral truths about which to agree.\(^{19}\) The best version of this argument, in my view, identifies moral philosophers, past and present—the supposed experts vis-à-vis moral inquiry—as the relevant class of disputants. Brian Leiter (2014) has recently advanced an argument of this kind that takes its inspiration from Nietzsche.\(^{20}\)

Leiter begins by describing the kind of disagreement that his metaphysical hypothesis—namely, that there are no moral truths—is meant to explain better than non-skeptical moral realism. He writes

> For what [Nietzsche] calls attention to is not “ordinary” or “folk” moral disagreement, but rather what seems to me the single most important and embarrassing fact about the history of moral theorizing by philosophers over the last two millennia: namely, that no rational consensus has been secured on *any* substantive, foundational proposition about morality. By a “foundational” moral proposition about morality, I am thinking of, for example, deontological or utilitarian theories which specify the criteria in virtue of which concrete or particular moral judgments are thought warranted: so e.g. “it is wrong to break this promise” is a concrete moral judgment, while “the wrong-making feature of an action is its effect on utility” is a foundational proposition. With regard to such foundational propositions, the history of moral philosophy is the history of intractable disagreement. Is the criterion of right action the reasons for which it is performed or the consequences it brings about? If the former, is it a matter of the reasons being universalizable, or that they arise from respect for duty, or something else? If the latter, is it the utility it produces or the

---

19 Mackie’s (1977) “argument from relativity” is the most famous version of this kind of argument.

20 In Leiter’s article, he is doing both Nietzsche interpretation and defending an argument from disagreement. Leiter attributes the argument I’ll consider in this section to Nietzsche, but I will attribute it and its supporting claims to Leiter, since Leiter is the one who does the most to defend them.
perfection it makes possible? If the former, is utility a matter of preference satisfaction (as the economists almost uniformly believe) or preference satisfaction under idealized circumstances…? (2014: 131)

The idea, then, is that moral philosophers have had roughly twenty-five hundred years to construct, and come to agreement about, a systematic account of ethics. Over that time, they have produced a host of incompatible philosophical theories: Aristotelianism, Kantianism, Humeanism, Utilitarianism, contractualism, and so on. Each purport to answer foundational questions in ethics, but none has been able to secure the endorsement of more than a fraction of moral philosophers. Thus, there is very little agreement and, moreover, very little in the way of rational persuasion among them. You almost never see a Kantian converting all at once (or even gradually) to Utilitarianism, or a Libertarian being rationally persuaded to embrace socialism, or any similar kind of conversion. For the most part, philosophers endorse (or incline toward) a certain ethical theory at the beginning of their careers and remain in that corner, as it were, forever. Their disagreements are thus amazingly intractable. Though philosophers often agree about particular moral judgments, that agreement ends as soon as they’re called on to explain why their particular judgments, and not their opponents’, are correct. And it is this fact, Leiter thinks—the fact that there is such intractable moral disagreement about foundational moral issues, not particular moral judgments—that calls out for explanation.

The explanation Leiter (inspired by Nietzsche) provides has two components: a psychological and a metaphysical component. He writes

[T]he “best explanation” argument asks: what is the best explanation for the fact that philosophical theories, in the sense just noted, reach different and quite intractable conclusions about foundational matters? Nietzsche’s skeptical answer will be that the best explanation is that the psychological needs of philosophers lead them to find compelling dialectical justifications for very different basic moral claims, and there are no objective moral facts to stand in the way of satisfying those psychological needs (2014: 134).

21 This is, at any rate, how Leiter sees things. For the sake of argument, I’m happy to accept his description of the sociology of moral philosophy.
The psychological component, then, is that philosophers have psychological needs that are satisfied when they believe their preferred moral theory. The reason there are so many different moral theories is that moral philosophers have many different psychological needs. Some moral theories will satisfy some philosophers’ psychological needs while other moral theories will satisfy others’.

The metaphysical component of Leiter’s explanation is that there are no moral truths to “stand in the way” of philosophers satisfying their psychological needs by believing their preferred moral theory. The thought seems to be that, if there were the kind of moral truths that non-skeptical realists posit, we would expect to see moral philosophers slowly but surely coming to believe them, despite the distorting influence of their psychological needs. We would therefore expect to see a slow but steady increase in the amount of agreement among moral philosophers as they discover the moral truths. Since we don’t see that kind of slow and steady increase in agreement, we can be confident that there are no such truths influencing their moral beliefs.

We might represent Leiter’s central argument as follows:

**The IBE Argument from the History of Moral Philosophy**

11. If there is long-standing, intractable disagreement among the (supposed) experts about a subject matter $S$, then the best explanation for the existence and character of that disagreement is that there are no facts about $S$.
12. There is long-standing, intractable disagreement among the (supposed) experts about ethics.
13. Therefore, the best explanation for the long-standing, intractable disagreement among moral philosophers is that there are no facts about ethics.

There are, of course, dangers associated with representing an inference to the best explanation as a deductive argument. For instance, Leiter might not accept (11) in full generality. But since Leiter accepts (11) as it pertains to the subjects that I will discuss here (i.e., ethics, epistemology, metaphysics), I can afford to simplify in this way.
Premise (12) says that ethics is characterized by long-standing, intractable disagreement among its supposed experts. But, of course, ethics is not unique in this respect. The same is true of, at least, metaphysics, epistemology, political philosophy, aesthetics, philosophy of language, philosophy of logic, and philosophy of science. If such disagreement in ethics is best explained by Leiter’s hypothesis that there are no moral facts, then a similar hypothesis should apply to the other subjects. That would entail that there are no facts about metaphysics, epistemology, philosophy of logic, and so on. Leiter anticipates this response—for metaphysics and epistemology, at least. He understands the challenge as an appeal to “companions in guilt.” Philosophers appeal to companions in guilt when they want to rebut an argument that has undesirable implications for their view. They do this by suggesting that the argument levelled against their preferred view overgeneralizes, having undesirable implications for views that their opponents do not wish to undermine. Leiter, however, is unmoved by appeals to companions in guilt. He writes,

Some recent writers (such as Bloomfield 2004 and Shafer-Landau 2005) think this kind of “companions in guilt” consideration counts in favor of moral realism, notwithstanding the disagreement among moral philosophers. It is not entirely clear why they rule out, however, the other natural conclusion. Nietzsche, as far as I can see, has no reason to resist it, since he believes that, as an explanatory matter, the moral commitments of the philosopher—at least the great philosopher—are primary when it comes to his metaphysics and epistemology (2014: 146).

Leiter’s (surprising) reply, then, is to admit that metaphysics and epistemology are just as “guilty” as ethics. On the Leiter/Nietzsche view, philosophers adopt their metaphysical and epistemological views because they support, or cohere with, their antecedently held ethical views. If so, then, on Leiter’s view, there are no facts of the matter when it comes to foundational questions in metaphysics and epistemology—just as in ethics.

But if there are no facts about metaphysics and epistemology, then what are we to make of Leiter’s premise (11)? (11) is either a metaphysical or an epistemological claim. Either way, if Leiter is correct that there are no facts of the matter when it comes to metaphysics and epistemology, then by
endorsing (11) and its implications, he has thereby undermined (11). He is committed to the claim that there is no fact of the matter about whether (11) is correct. But if that is so, then we clearly should not accept his IBE argument from the history of moral philosophy. For, by its own standards, there is no fact of the matter about whether its first premise, (11), is correct.

5. Arguments from Disagreement in Ideal Conditions

The final popular form of argument I’ll consider begins with the supposition that there is significant fundamental moral disagreement among people generally—both philosophers and non-philosophers. Fundamental moral disagreement, as it’s meant here, is moral disagreement that would persist even between agents who are fully-informed about non-moral matters and are ideally-rational. The argument then proceeds to draw a troubling metaphysical conclusion about morality, namely, that, for many debates in ethics, there is no fact of the matter about which view is correct. And if so, then our understanding of morality is deeply flawed and non-skeptical realism is false. After all, on non-skeptical realism, we do not have a deeply flawed understanding of morality. The best version of this argument comes from Richard Rowland (2016).

Rowland begins by arguing for an accessibility condition on moral truths. He thinks that, if there are moral truths, they must be epistemically accessible to (i.e., discoverable by) some possible agent. He reasons as follows. If there is a fact of the matter about the moral status of some action Φ, then that fact makes a difference to how some possible agent ought to act. For example, if it’s (im)permissible to eat meat, then that fact makes a difference to how we ought to live. But if it is impossible for any possible agent to discover the truth about the moral (im)permissibility of eating meat, then that fact wouldn’t make a difference to how any possible agent ought to live. So, there wouldn’t be any fact of the matter about whether it’s permissible to eat meat. Thus, the only way that
an action can have a moral status is if some possible agent could discover the truth about its moral status. I’ll call this the *accessibility condition*.

Rowland then argues that, for many debates in moral philosophy, it’s not unlikely that the putative moral status of the relevant act is inaccessible to any possible agent. For example, he thinks that it’s not unlikely that the putative facts about the moral status of vegetarianism, bombing Syria, torture, the death penalty, breaking promises, pushing the fat man off the bridge, and lying are inaccessible to any possible agent because disagreement about the moral status of these actions would persist even among agents in epistemically ideal conditions (2016: 2). That is, even if the disputants in these debates were given full non-moral information and were endowed with perfect reasoning capacities, they would still disagree about the moral status of these actions. If so, Rowland argues, then the supposed facts about the moral status of these actions would be inaccessible to any possible agent. So, given the accessibility condition, there would be no fact of the matter about the moral status of these actions. Our understanding of morality would thus be deeply flawed and non-skeptical moral realism would be false. We could represent the argument as follows:

*The Argument from Disagreement in Ideal Conditions*

14. For any action \( \Phi \), if \( \Phi \)-ing has a moral status, then the moral status of \( \Phi \)-ing is accessible to some possible agent.
15. For many debates in ethics, the moral status of \( \Phi \)-ing is inaccessible to any possible agent.
16. So, for many debates in ethics, \( \Phi \)-ing has no moral status.
17. If (16), then our understanding of morality is deeply flawed and non-skeptical realism is false.
18. Therefore, our understanding of morality is deeply flawed and non-skeptical realism is false.

I’ll focus on (14)-(16), since I accept (17). (14) is the accessibility condition, which I’ll leave unchallenged. The support for (15) includes two suppositions. The first is that there is fundamental disagreement about many debates in moral philosophy. I’m happy to grant that. The second is that, if
there is fundamental disagreement about the moral status of Φ-ing, then the moral status of Φ-ing is inaccessible to any possible agent. This is where I will focus my criticism.

Rowland doesn’t give an account of what epistemic accessibility consists in, but he does offer a necessary condition for a proposition’s being epistemically accessible. It’s a condition that he believes is not met for many propositions that are the subject of fundamental disagreement in ethics. Here’s the condition.

(2.1) A has epistemic access to p at T1 only if there is a justification R that A can justifiably believe p on the basis of at T1 that is such that A’s belief that p on the basis of R would not be defeated in more ideal or idealized conditions (2016: 12).

In other words, an agent has epistemic access to a proposition p only if she can justifiably believe p in epistemically ideal conditions. The key claim in Rowland’s case for the claim that ideally informed, ideally rational agents engaged in a moral disagreement do not meet this condition with respect to many moral propositions is

(2.3) If idealized reasoners A and B hold conflicting beliefs about the moral status of Φ-ing, know this, and know that they are idealized reasoners, then neither A nor B are justified in holding conflicting beliefs about the moral status of Φ-ing; A and B are only justified in suspending belief about the moral status of Φ-ing (2016: 12).

The idea is that, if A and B are idealized reasoners and know this about one another, then they each know of the other that they have the same evidence (that is, all the non-moral facts) and have made all and only correct steps in reasoning. So, neither can claim to have more (or different) information, and neither can claim to have followed a better line of reasoning than the other. Rowland thus concludes that neither can claim to have the “better” view, and neither would be justified in holding their own view in the face of this disagreement. They’re both rationally required to suspend judgment about the disputed proposition.
Though Rowland formulates (2.3) in terms of moral disagreement, there’s no motivation for restricting it to moral, but not normative, disagreement more generally. Extending (2.3) in this way yields

\[(2.3*)\] If idealized reasoners \(A\) and \(B\) hold conflicting beliefs about the normative status of \(\Phi\)-ing, know this, and know that they are idealized reasoners, then neither \(A\) nor \(B\) are justified in holding conflicting beliefs about the normative status of \(\Phi\)-ing; \(A\) and \(B\) are only justified in suspending belief about the normative status of \(\Phi\)-ing.

\(\Phi\)-ing would now refer, not only to acting in various ways, but also believing in various ways.

It’s clear that (2.3*) is itself a normative claim, since it’s about the conditions under which holding a certain belief is not epistemically justified. The first question we should ask, then, is: Would (2.3*) itself be the subject of disagreement among agents in epistemically ideal conditions? If so, then, by (2.3*), even agents in ideal conditions would not be justified in believing (2.3*). So, by (2.1), no possible agent would have epistemic access to (2.3*). Combined with premise (14), this would entail that (2.3*) has no normative status—i.e., there would be no fact of the matter about whether (2.3*) is correct.

Obviously, then, Rowland must insist that (2.3*) would not be the subject of disagreement in ideal conditions. But what could justify this insistence? Why should anyone believe that agents in epistemically ideal conditions would come to agreement about (2.3*) and, furthermore, agree to accept, rather than to reject or suspend judgment about, (2.3*)? I can imagine two strategies for justifying this claim. The first is to endorse Conciliationism. The second is to endorse the Interpersonal Uniqueness thesis. I’ll consider each strategy in turn.

If Conciliationism is correct, then that would explain why ideally informed, ideally rational agents would suspend judgment about controversial moral propositions once they become aware that their ideally informed, ideally rational interlocutor disagrees with them. Plausibly, such agents are epistemic peers and Conciliationism requires disagreeing epistemic peers to suspend judgment about
any proposition that is under dispute between them. There are, however, two problems with appealing to Conciliationism to rule out the possibility that (2.3*) would itself be the subject of disagreement in ideal conditions. The first, as I argued in section 2, is that, even by conciliationists’ lights, we cannot now rationally believe that Conciliationism is correct, given the current controversy about it among excellent philosophers. Thus, we cannot now rationally accept Rowland’s argument if it depends on Conciliationism. I’ll discuss the second problem with appealing to Conciliationism to justify (2.3*) in a moment.

The second way to justify the claim that (2.3*) would not be the subject of disagreement in idealized conditions is to appeal to

**The Interpersonal Uniqueness Thesis (IUT):** Given a body of evidence \( E \), a proposition \( p \), and any two agents \( A \) and \( B \), there is precisely one doxastic attitude toward \( p \) that is rationally permissible for \( A \) and \( B \) relative to \( E \).

If this were correct, then ideally rational agents with the same evidence would never disagree about anything. But, of course, the IUT is highly controversial, too.\(^{22}\) Even now, a debate is raging between permissivists who deny the IUT and non-permissivists who accept it.\(^{23}\) This debate is at an early stage in its development, so it’s not clear how it will turn out. So, it’s certainly not clear that the consensus, should it ever emerge, will settle on acceptance (rather than agnosticism about or rejection) of the IUT.

Indeed, the IUT is at least as controversial as any of the moral issues about which Rowland suggests there might be fundamental disagreement. Speaking of the IUT, Matt Kopec and Michael Titelbaum write

\[T\]he question of whether Uniqueness is true has many of the trappings of a classic philosophical puzzle. The thesis seems obviously false to many philosophers, and obviously true to many others. The arguments in favor of Uniqueness are at first rather

\(^{22}\) Enoch makes this point in (2011: 208).
compelling, but a closer inspection reveals various flaws. Nevertheless, those who think Uniqueness is false often find it difficult to devise any counterexamples that are able to convert the Uniqueness defenders (2016: 189).

If this is correct, then the IUT, like Conciliationism, shares all the relevant features of debates about vegetarianism, bombing Syria, torture, and so on—debates about which Rowland suggests there might be fundamental disagreement. Thus, any reason to believe that moral disagreement would persist among ideal reasoners about these issues favors believing that epistemic disagreement would persist among them, too—especially about the IUT and Conciliationism. But if such disagreement would persist, then, by Rowland’s own lights, these principles are not epistemically accessible to any possible agent. Combined with (14)—the accessibility condition—this entails that there is no fact of the matter about whether these epistemic principles are correct. If so, then Rowland has undermined the most promising support for (2.3*)—his central claim in his argument from disagreement. Moreover, since the support for (2.3*) likely depends on the IUT or Conciliationism, and these principles would be the subject of disagreement in ideal conditions, there is excellent reason to believe that (2.3*) would itself be the subject of disagreement in ideal conditions. By Rowland’s own lights, that entails that there is no fact of the matter about whether (2.3*) itself is correct. We therefore shouldn’t accept his argument from disagreement, which crucially depends on (2.3*).

6. Conclusion

I’ve now considered four different popular forms of argument from disagreement and argued that each fail in their own way. It’s clear, however, that the flaw in these arguments is structural—each fails for roughly the same underlying reason. Each argument attempts to support a very strong conclusion about the metaphysics or epistemology of morality. To do this, these arguments must employ very strong metaphysical or epistemological premises. But given the pervasiveness of disagreement in philosophy—especially about metaphysics and epistemology—such strong premises
are always, or nearly always, the subject of deep, widespread, intractable disagreement. If such disagreement is sufficient to support troubling conclusions about morality, as these arguments purport to show, then such disagreement is sufficient to support troubling conclusions about metaphysics and epistemology, too. By these argument’s own standards, then, such disagreement undermines the strong metaphysical and epistemological premises these arguments employ to derive their radical conclusions about morality. In short, then, given the pervasiveness of disagreement in philosophy, these arguments are destined to undermine themselves.
References


Sampson, E. ms. “Can We Rationally Believe Conciliationism?”


