What Pessimism about Moral Deference Means for Disagreement

James Fritz

This is a post-peer-review, pre-copyedit version of an article published in Ethical Theory and Moral Practice. The final authenticated version is available online at: http://dx.doi.org/*10.1007/s10677-017-9860-8*

**Abstract** Many writers have recently argued that there is something distinctively problematic about sustaining moral beliefs on the basis of others’ moral views. Call this claim *pessimism about moral deference*. Pessimism about moral deference, if true, seems to provide an attractive way to argue for a bold conclusion about moral disagreement: moral disagreement generally does not require belief revision. Call this claim *steadfastness about moral disagreement*. Perhaps the most prominent recent discussion of the connection between moral deference and moral disagreement, due to Alison Hills, uses pessimism about the former to argue for steadfastness about the latter. This paper reveals that this line of thinking, and others like it, are unsuccessful. There is no way to argue from a compelling version of pessimism about moral deference to the conclusion of steadfastness about moral disagreement. The most plausible versions of pessimism about moral deference have only very limited implications for moral disagreement.

Many find it plausible that there is something distinctively problematic about moral deference. Call this thesis *pessimism about moral deference* (or, for short, *pessimism*).[[1]](#footnote-1) Pessimists come in many shapes and sizes, but most do not argue that relying on others’ moral views is always forbidden. They seek, instead, to identify the most troublesome varieties of moral deference and explain what makes those varieties troublesome.

Pessimism about moral deference seems, at first, to have clear connections to a different topic within social epistemology: responsiveness to disagreement. Say that, in a range of cases, moral beliefs based on others’ views are problematic. This suggests that, within that range of cases, I should not give others’ views much weight in my moral thinking. Perhaps, for the same reasons, I should (in the relevant cases) not allow the fact that others disagree with me to change my level of confidence in my moral beliefs.[[2]](#footnote-2) Call this conclusion *steadfastness* *about moral disagreement*.

It can seem patently obvious that there is a close relationship between pessimism about moral deference and steadfastness about moral disagreement. One prominent pessimist, Alison Hills, even offers a brief argument from the former to the latter (2010). But upon closer reflection, the relationship between pessimism about moral deference and steadfastness about moral disagreement is not so simple. Pessimists offer a wide variety of explanations of the distinctive demerit of moral deference, and different explanations have different implications for disagreement.

This paper argues that there is no way to argue from a compelling version of pessimism about moral deference to the conclusion of steadfastness about moral disagreement. Some brands of pessimism do imply that a great deal of steadfastness is appropriate, but those brands are independently unattractive. By contrast, more prominent, plausible versions of pessimism have only very limited implications for moral disagreement.

1. Deference and Disagreement

This section sets the stage for the discussion to come by illustrating a difference between moral deference and responsiveness to moral disagreement. The difference is this: moral deference always involves a person coming to sustain a moral belief on the basis of another person’s (apparent) beliefs.[[3]](#footnote-3) Responsiveness to moral disagreement, by contrast, need not have this feature. In some cases, a person is responsive to moral disagreement without coming to sustain any moral belief on the basis of another person’s.

Just what is moral deference? Paradigmatically, my deferring to you about some moral matter involves my both *forming* and *sustaining* some moral belief on the basis of your (apparent) moral belief. By deferring to you, in other words, I come to have a moral belief that I did not have before—a belief that, loosely speaking, is based “on the say-so of others” (Hills 2009, 94).

It’s worth noting that some genuine cases of moral deference may not fit this paradigm. Suppose, for instance, that I believe that eating factory-farmed meat is morally wrong. I then talk to a professional ethicist who agrees. I am so dazzled by her professionalism and sophistication that I entirely lose track of the reasons for which I used to consider eating factory-farmed meat morally wrong; indeed, I come to sustain my already-held belief that eating factory-farmed meat is wrong *solely* on the basis of the fact that she believes it. In this case, I do not form any new moral belief at all. But I do replace the basisfor an extant moral belief with a new basis. Perhaps this sort of response to testimony also deserves to be called ‘moral deference’. If so, then the gloss above is not quite accurate; some moral deference does not involve the formation of a new moral belief. In order to leave open that possibility, my discussion below will not assume that deference involves the formation of new beliefs; I will focus, instead, on the fact that moral deference involves coming to sustain a moral belief on the basis of others’ (apparent) moral beliefs.

We are now ready to note a crucial difference between deference and responsiveness to disagreement. Consider the difference between the following three reactions to disagreement:

**Reduce Confidence** *P* is a moral proposition. At *t0*, you believe that –*p.* You then learn that I believe that *p*. In response, you become less confident that *–p*. At *t1*, you still believe that *–p*, but you do so less confidently than you did before.

**Suspend Judgment** *P* is a moral proposition. At *t0*, you believe that –*p.* You then learn that I believe that *p*. In response, you abandon your belief that *–p*. At *t1*, you neither believe that *p* nor believe that *–p*; you suspend judgment as to whether *p*.

**Switch Beliefs** *P* is a moral proposition. At *t0*, you believe that –*p.* You then learn that I believe that *p*. In response, you abandon your belief that *–p* and form the belief that *p*. At *t1*, you believe that *p* and sustain this belief on the basis of mine.

These responses to disagreement are arranged in ascending order of the degree to which you alter your doxastic state. When confronted with disagreement, you might reduce your confidence but retain your belief, you might suspend judgment, or you might abandon your initial belief and adopt its opposite.

The core question of this paper is: which responses to moral disagreement, if any, are ruled out by pessimism about moral deference? To begin answering this question, note a striking trait of **Switch Beliefs**. **Switch Beliefs** *just is* a case of moral deference as glossed above; when you switch beliefs, you come to sustain one of your beliefs on the basis of someone else’s (apparent) moral belief. Now, granted, most cases of moral deference described in the current literature are not stipulated to start from the point of disagreement with an interlocutor. But most leave that possibility open. Moreover, to the extent that pessimists consider deference problematic, they are likely to consider **Switch Beliefs** problematic in just the same way. **Switch Beliefs,** then, is morally deferential behavior of just the sort that concerns contemporary pessimists.

The fact that **Switch Beliefs** is a case of moral deference tells us something about the relationship between moral deference and moral disagreement. We should expect pessimist views of moral deference to find at least some responses to moral disagreement distinctively problematic. On pessimist views, the best sort of moral thinker will generally not respond to moral disagreement by abandoning her contested belief, adopting her opponents’ contrary belief, and sustaining it on the basis of their moral view.

On the other hand, pessimists about moral deference are not obviously committed to any particular conclusion about the appropriateness of reducing confidence or suspending judgment in response to disagreement. Perhaps, even if switching beliefs (and thereby deferring) in response to disagreement is bad, these more moderate responses are entirely acceptable. To see this, consider a simple and wholly implausible version of pessimism about moral deference. On this toy view, moral deference is bad because it is morally impermissible for a person to adopt or continue holding any moral belief *p* whenever she is aware that any other person believes that *p*. This principle rules out certain ways of forming and continuing to hold moral beliefs, but it does not rule out any other sorts of doxastic behavior. Notably, though it forbids moral deference, it does not forbid reducing confidence or suspending judgment in response to disagreement. Other pessimist views that locate the badness of moral deference in the way that we adopt or sustain outright moral beliefs will tend to have similarly limited implications for moral disagreement.

Sections 3 and 4 of this paper show that the most plausible contemporary defenses of pessimism about moral deference are, in this respect, like the toy example just considered. Though they can explain what would be bad about responding to disagreement by switching beliefs, they cannot explain what would be bad about reducing confidence or suspending judgment. The most plausible forms of pessimism about moral deference, then, do not entail any striking conclusions about the appropriate response to moral disagreement.

But not all brands of pessimism about moral deference have such limited implications about disagreement. Section 2 considers some initially tempting, but ultimately implausible, ways of developing pessimism about moral deference. Several of these views have quite sweeping implications for moral disagreement. In that respect, they provide a useful contrast to the more plausible views considered in sections 3 and 4.

1. Unattractive Versions of Pessimism

This section surveys several unattractive approaches to pessimism about moral deference. It mentions some problems for these brands of pessimism, and it notes their implications for disagreement.

* 1. Accessibility

It’s tempting to suppose that, in some important sense, moral truths are accessible to us all. Perhaps this is what is wrong with moral deference; when we defer, we show that we are unable or unwilling to do the work of any minimally competent inquirer into moral truth.

To the extent that this claim about accessibility is true, it seems to support not only pessimism but also steadfastness about moral disagreement. Most views about the epistemology of disagreement take pains to establish that we can justifiably remain confident when a stranger denies some obvious truth.[[4]](#footnote-4) So, to the extent that moral truths are obvious, steadfast belief in those truths in the face of disagreement is likely appropriate.

This strategy, however, relies on an implausible assumption: that moral truths are, in a way that sets them apart from non-moral truths, generally accessible to competent thinkers. Sarah McGrath (2009, 2011) offers a dilemma for the pessimist who relies on this claim. On one reading of “accessible”, it’s implausible that moral truths are generally accessible to all competent thinkers; the extent of disagreement suggests that many sincere, committed moral thinkers are unable to reason their way to many important moral truths without help. On other readings of “accessible,” it may be true that moral truths are accessible to all competent thinkers, but this observation does not provide support for pessimism about moral deference. Consider, by way of comparison, very complicated addition problems. It’s true in *some* sense that the solutions to these problems are accessible to all competent inquirers into addition. But this does not make us suspicious of deference about complicated addition problems.[[5]](#footnote-5)

So there are good reasons to be suspicious of the claim that moral truths are, in a distinctive way, accessible to competent thinkers. This brand of pessimism might have sweeping implications for moral disagreement, but pessimists have more attractive options elsewhere.

* 1. Reliability

One approach to pessimism proceeds through observations about the reliability of other moral thinkers. The pessimist might suggest that one should only defer to those whom one has good reason to consider reliable and argue that we seldom have good reason to consider another person’s moral thinking reliable.[[6]](#footnote-6) More modestly, she might claim that one should only defer to those whom one has good reason to consider *more reliable than oneself* and argue that we seldom have good reason to consider others more morally reliable than we are.

To the extent that they are plausible, these observations justify quite robust steadfastness about moral disagreement. Even the most conciliatory views about the epistemology of disagreement acknowledge that disagreement about *x* does not require reduced confidence unless we have good reason to think that our interlocutors are reliable about *x*.[[7]](#footnote-7) If we never have such reason, then disagreement never requires us to lower our confidence, much less suspend judgment.

But observations about reliability are not a persuasive way to motivate pessimism about moral deference. There are a host of problems with this strategy,[[8]](#footnote-8) but perhaps the most serious problem comes from examples of deference that, although they might seem problematic to pessimists, involve interlocutors who are known to be very reliable.

David Enoch (2014, 230) offers a convincing example of this sort. In the example, Enoch’s judgment about military engagement is easily swayed by emotion, but his colleague’s judgment is more careful. On many past occasions, although they disagreed about the permissibility of specific military actions at first, Enoch has come to agree with his colleague over time. Here, Enoch is surely justified in believing that his colleague’s moral judgment about war is more reliable than his own. Many pessimists will argue that there would be some problem with Enoch deferring to his colleague.[[9]](#footnote-9) But that problem cannot be the reliability of Enoch’s colleague, or Enoch’s awareness of that reliability.

Pessimists about moral deference, then, should not support their view through worries about reliability. Fittingly, contemporary pessimists have avoided this strategy.

* 1. Autonomy and Authenticity

Some have been tempted by the thought that a person’s moral thinking should be *autonomous* or *authentic*. At first glance, these thoughts support both pessimism about moral deference and steadfastness about moral disagreement; aspiring to autonomy or authenticity suggests striving for independence from others’ moral beliefs. But there are issues with this approach to pessimism about moral deference. Most importantly, there may not be a notion of autonomy or authenticity that is both valuable and incompatible with deference and conciliation.

Some straightforward ways of developing these notions result in an extensionally inadequate pessimism. On one approach, for instance, moral thinking is *autonomous* to the extent that the thinker freely decides to adopt his moral beliefs. But this seems entirely compatible with paradigmatic deference; to the extent that I can freely decide to adopt any belief, I can freely decide to adopt someone else’s belief (McConnell 1984, 210-11). On one approach to authenticity, my beliefs are authentic just to the extent that their contents *seem* true to me (cf. Pasnau 2015, 2326-7). But even a belief that is authentic in this sense are apparently objectionable in just the same way that paradigmatic cases of moral deference are apparently objectionable. Say that whenever my pastor asserts a moral proposition, that proposition comes to seem obviously, primitively true to me, and I come to believe it. Even if this pattern of belief-formation is authentic, it is no less troubling than paradigmatic cases of deference. A version of pessimism that appeals to authenticity, but cannot rule out this sort of deference, is extensionally inadequate.

Some ways of developing the notions of autonomy and authenticity do not obviously share these problems with extensional adequacy. Pessimists might suggest, for instance, that a person’s moral thinking is authentic to the extent that it gives “expression to the [person’s] true self” (Mogensen 2015, 17). But these approaches, even if they get the right results about cases, seem to push the task of justifying pessimism about deference back to different questions: for instance, is there a “true self” such that it really is important to express that self in one’s beliefs, and deference inhibits that expression?[[10]](#footnote-10)

Sections 3 and 4 consider views that attempt to answer the deeper question here. They characterize the way in which we ought to form our moral views, and they also explain just what would be so valuable about forming our views in that way. Compared to such projects, barefaced appeals to autonomy or authenticity can seem explanatorily shallow. Perhaps that is why such appeals have been so rare (and, when offered, so tentative) in the contemporary literature about moral deference.[[11]](#footnote-11)

We’ve now considered several lines of thought that seem apt to support both pessimism about moral deference and steadfastness about moral disagreement. But, as we’ve seen, there are good reasons not to embrace these approaches to pessimism. The remainder of the paper considers more compelling versions of pessimsim. These views, we’ll discover, have only very limited implications for moral disagreement.

1. Understanding

By far the most prominent brand of pessimism about moral deference involves appeals to the value of moral understanding. For brevity, I’ll call this sort of view *understanding-pessimism*. Early defenses of understanding-pessimism can be found in Nickel (2001) and Hopkins (2007). Sarah McGrath (2011) has also offered a tentative endorsement. But the standard-bearer for understanding-pessimism is Alison Hills. Hills is also one of the only writers to draw an explicit connection between understanding-pessimism and steadfastness about moral disagreement. In her (2010), Hills offers a brief argument to the effect that suspending judgment in response to disagreement is problematic for just the same reasons that moral deference is problematic.

This section offers a brief overview of Hills’s understanding-pessimism. It then shows that, even if Hills’s view about deference is true, it does not vindicate her conclusion about moral disagreement. Like any pessimist view, understanding-pessimism can cast doubt on the practice of responding to moral disagreement by switching beliefs. But understanding-pessimism does not imply that there is any problem with more moderate responses to disagreement.

3.1 Hills on Deference and Disagreement

Hills characterizes the state of understanding why *p* as a complex, factive propositional attitude. In order to understand why *p*, you must believe that *p*. You must also believe that *q* is why *p*. Importantly, you must also go beyond merely believing that *q* is the reason why *p*; you must *grasp* the reasons why *p* is true. According to Hills, fully grasping reasons requires the abilities to:

1. follow an explanation of why *p* given by someone else;
2. explain why *p* in your own words;
3. draw the conclusion that *p* (or that probably *p*) from the information that *q*;
4. draw the conclusion that *p*’ (or that probably *p*’) from the information that *q*’ (where *p’* and *q’* are similar to but not identical to *p* and *q*);
5. given the information that *p*, give the right explanation, *q*;
6. given the information that *p’*, give the right explanation, *q’*. (2009 ,102-3)[[12]](#footnote-12)

To have *moral* understanding, on Hills’s account, is just to understand why *p* for some moral proposition *p*.

Moral understanding, Hills claims, is a distinctly valuablestate. For one, it tends to make a person a reliable moral judge (2013, 555). Moral understanding is also a crucial ingredient in virtue and morally worthy action; it allows agents, in their thought and their action, to respond directly to moral reasons (2009, 110-4; cf. 2010, 198-213; 2015, 12). The same cannot be said for mere true moral belief, or even for moral knowledge.

The fact that moral understanding is an appropriate aim for moral thinkers, Hills proposes, vindicates pessimism about moral deference. Suppose a person is “perfectly capable of working out for herself” (2010, 219) whether lying is morally bad. But, rather than going to the trouble of working out the question, she asks her pastor and simply takes his word on the subject. Even if she thereby comes to know that lying is morally bad, Hills argues, there is something problematic about the way in which this person is conducting her doxastic life. A mature moral agent should strive for moral understanding, and to gain moral understanding, she must base her moral beliefs on the reasons for which they are true (2010, 192). Our imagined inquirer does not strive to gain this robust basis for her moral beliefs—rather, she adopts the “rival basis” (2010, 220) of her pastor’s testimony. Here is what Hills has to say about this “rival basis” for moral belief:

At best it is worthless, since you need to base your belief on another ground anyway, and at worst it is harmful, since it could make it less likely that your belief is grounded in the right way… .(2010, 220)

In short, the idea is this: moral inquiry should aim at moral understanding. And moral deference, even when it issues in knowledge, is at best an idle step in the search for moral understanding, because it cannot provide the basis for moral belief that is required to grasp the reasons for one’s belief. Therefore, the best sort of moral inquiry does not involve moral deference.

Hills’s pessimism about deference has limits. She acknowledges that, in certain situations, “it may be better to trust the judgment of others” (2009, 125; cf. 2010, 229-30) than to continue the search for understanding. For instance, one might have good reason to suppose that moral understanding is out of reach, but that moral knowledge can be gained through deference. But, Hills notes, it is far from clear just how often situations of this sort arise (2009, 125). This explains why we are generally, and rightly, suspicious of moral deference; it is usually best understood as a failure to adequately pursue moral understanding.

Suppose that understanding-pessimism correctly explains the problematic nature of moral deference: deference often manifests a failure to pursue moral understanding. What would this mean for responses to moral disagreement? Hills argues that certain ways of responding to disagreement, like paradigmatic instances of moral deference, manifest a failure to aim at moral understanding. She calls attention to a difference between two ways in which a person might revise his moral belief that *p* in response to a disagreement.

On one hand, he might, in the course of disagreement, learn of new arguments or evidence that bear directly on whether *p.* A disagreement about whether to eat factory-farmed food, for instance, might bring to someone’s attention new data about how animals are treated in factories. Hills calls considerations of this sort “*explanatory evidence*… [that is, evidence] that could be used to explain why [the target proposition] is true” (2010, 222). Hills acknowledges that learning new explanatory evidence in the course of disagreement can bring us to revise our beliefs in an entirely respectable way (2010, 221). After all, on Hills’s view, moral understanding must be based on a grasp of explanatory evidence. Responsiveness to such evidence should certainly be part of any sincere, effective search for moral understanding.

On the other hand, Hills considers it problematic for a person to revise a moral belief in response to the *mere fact* that someone else disagrees. The fact that someone else believes not-*p* is *non-explanatory evidence* for not-*p*; though it can provide “significant and relevant evidence” in favor of not-*p*, it “could not be used to explain why not-*p* is true” (2010, 222). According to Hills, it is inappropriate to revise one’s moral beliefs in response to non-explanatory evidence. So the right kind of inquirer into morality does not suspend judgment about a moral matter in response to the mere fact of disagreement.

What would be bad about revising one’s moral beliefs on the basis of non-explanatory evidence? Here, Hills appeals to the importance of moral understanding. The aim of moral belief is moral understanding (2010, 227-230; 2015, 14-16), which is important (among other reasons) because, without it, we cannot act for the right reasons. Hills connects these observations as follows:

…if you give weight to [non-explanatory evidence], you are not forming your moral belief for the right reasons. The structure of your moral beliefs is not mirroring morality, and when you act you cannot act for the right reasons… .(2010, 222)

In this passage, Hills gestures toward a unified explanation for pessimism about moral deference and steadfastness about moral disagreement. When we revise our beliefs based on disagreement, as when we defer, we are revising our beliefs on the basis of non-explanatory evidence. But someone who aspires to moral understanding should neither form nor revise beliefs in this way; rather, she should make her moral thinking solely responsive to the reasons that make moral propositions true. Only in this way, Hills claims, can her moral thinking put her in a position to reap the benefits of understanding.

3.2 Why Understanding-Pessimism does not Justify Steadfastness

It’s perhaps easiest to see the problem with Hills’s argument by considering a very weak way of responding to disagreement. Let’s imagine a contrast between two moral thinkers, Ephraim and Wanda. Both form all their moral beliefs solely on the basis of careful thinking about explanatory evidence. But Wanda’s moral thinking, unlike Ephraim’s, also conforms to an additional principle, which I’ll call Minimal Humility:[[13]](#footnote-13)

**Minimal Humility** Suppose that I believe moral proposition *p*. Suppose that I then learn a fact about the people who have spent more time than I have considering explanatory evidence for and against *p*: I learn that the vast majority of such people believe that not-*p*. In such circumstances, and as long as a slight reduction in my confidence that *p* would not involve my suspending judgment about *p*, I should become slightly less confident that *p*.

To follow Minimal Humility is to revise one’s beliefs on the basis of non-explanatory evidence; the opinions about *p* of other serious thinkers could not be used to explain why *p* is false. This is precisely the problem that Hills cites with certain ways of responding to moral disagreement. So, if Hills’s thinking about non-explanatory evidence is right, Wanda should be worse-off than Ephraim with respect to the pursuit of moral understanding.

But Wanda is just as well-placed as Ephraim to acquire moral understanding and the benefits that attend it. For any given moral proposition *p* such that Ephraim understands why *p*, Wanda can also understand why *p*. She can retain outright belief that *p,* and she can retain outright belief that *q* is why *p*. She can grasp the reasons that make *p* true, and she can retain all the abilities enumerated by Hills; she can explain why *p* in her own words, she can apply the relevant moral reasoning to other nearby cases, and so on. To the extent that Ephraim’s actions are performed for the right reasons, and to the extent that his actions can have moral worth, the same can be said of Wanda’s actions. Wanda’s adherence to Minimal Humility, though it makes her beliefs responsive to non-explanatory evidence, does not frustrate her pursuit of moral understanding in any way. So if understanding-pessimism reveals some problem with revising one’s moral beliefs in response to disagreement, that problem cannot be responsiveness to non-explanatory evidence.

The understanding-pessimist, then, cannot convincingly argue that the pursuit of moral understanding precludes responding to moral disagreement by lowering one’s confidence. But perhaps a weaker position is available: though the mere fact of moral disagreement can rationalize reductions in confidence, it can never rationalize suspending judgment about a moral proposition.[[14]](#footnote-14) After all, when I suspend judgment as to *p*, I lose my belief that *p* and thereby my understanding that *p*. So suspense of judgment presents a *prima facie* threat to moral understanding in a way that reduced confidence does not. Perhaps this straightforward insight can explain why understanding-pessimists should advocate for steadfast belief in the face of moral disagreement.

This strategy will not work either. To see this, first note that an understanding-pessimist must allow that it is sometimes permissible to suspend judgment about some moral questions for reasons that have nothing to do with non-explanatory evidence. Some moral questions are very difficult to answer, and sometimes we simply do not have the experiences or non-moral information to justify outright belief. Sometimes we must reflect on complicated arguments before we would be justified to draw a moral conclusion, and suspending judgment is the only reasonable stance in the interim. Next, note that this sort of behavior poses just the same *prima facie* threat to moral understanding that suspending judgment in response to disagreement poses. Whether I suspend judgment about *p* in order to give myself more time to reflect or in response to worrisome non-explanatory evidence, I thereby give up the possibility of understanding why *p*. So why is suspending judgment in response to moral disagreement particularly problematic? The answer cannot be that suspending judgment about *p* precludes understanding why *p*; this is no objection at all to suspense of judgment in the course of everyday inquiry. Nor can the answer be that suspending judgment in response to disagreement involves responding to non-explanatory evidence; as the example of Minimal Humility shows, the understanding-pessimist has no principled objection to responsiveness to non-explanatory evidence.

The understanding-pessimist might attempt to renew her argument by reasoning as follows. Suspending judgment about *p* in the course of everyday reflection is part of a pattern of doxastic behavior that, under favorable conditions, tends to result in moral understanding. Suspending judgment about *p* in response to *non-explanatory* evidence, by contrast, manifests a disposition that does not tend to provide moral understanding under favorable conditions.

This reply fails. Suspending judgment in response to disagreement can, in principle, play a crucial role in a scheme of doxastic behavior that provides moral understanding. To see why, first note that any instance of suspended judgment never suffices, by itself, to provides moral understanding. Just what role, then, does suspense of judgment play in the search for moral understanding? Plausibly, it helps us avoid falsehood; loosely speaking, it clears the way for further inquiry that (under favorable conditions) will lead to understanding of the truth. But there is no reason to suppose that it can only play this role when it is a response to explanatory evidence, rather than non-explanatory evidence like disagreement. Both sorts of evidence can help to point us away from falsehood, and they can help us clear the way for understanding of the truth.

This section has, so far, had fairly modest aspirations. It has shown that Hills’s particular proposal connecting understanding-pessimism to the epistemology of moral disagreement is unsuccessful. I’ll close by suggesting that the problem with her proposal generalizes to understanding-pessimists of all stripes; a focus on the importance of moral understanding is better-suited to show problems with ways of *forming and sustaining* beliefs than with ways of losing confidence in, or abandoning, those beliefs. This means that, even if the understanding-pessimist can convincingly cast doubt on moral deference, her view will have limited implications for disagreement.

3.3 The Twin Aims of Understanding

The core proposal of understanding-pessimism is that moral belief aims not solely at truth or knowledge, but at understanding. But surely, the understanding-pessimist should not say that the *only* aim of moral belief is understanding of true moral propositions. She must also acknowledge that moral beliefs aim to avoid merely apparent understanding of false moral propositions.[[15]](#footnote-15) Without this second goal, moral belief’s aim would be fully satisfied by the formation of an outright belief in every moral proposition (including every moral proposition of the form q *is why* p) and the acquisition of a suite of related abilities.

Now, there are two ways in which an epistemic norm might take into account the importance of avoiding merely apparent understanding of falsehood. First, it might make explicit recommendations about what to do when one’s beliefs are false. For example, consider a norm like the following: when *p* is false, abandon your beliefs in *p* in the face of disagreement, but when *p* is true, remain confident in *p* in the face of disagreement. This first sort of norm, though it may recommend steadfast belief in some cases of disagreement, does not recommend nearly as much as Hills does. That’s because, when my moral belief is false, it recommends that I revise that belief in response to mere non-explanatory evidence.

Second, an epistemic norm might be framed as the sort of advice that, though fallible, is likely to reduce belief in falsehood under favorable conditions and in the long run. Such a norm might advise, for instance, revising or abandoning a belief in *p* when someone who seems very reliable claims that *p* is false. This second sort of norm, unlike the first sort, seems potentially helpful as advice. It also departs from the first sort in that it does not aim at exceptionless success; in certain misleading evidential conditions, this norm might point away from true belief. Like the first sort of norm, however, this second sort of norm falls short of the steadfastness that Hills defends. It acknowledges the importance of responding to adequately powerful non-explanatory evidence, like disagreement, in the avoidance of false moral belief.

On either construal of the epistemic norms governing moral disagreement, then, it is entirely intelligible that sincere pursuit of the twin aims of moral understanding would involve belief revision in response to disagreement. Perhaps the same cannot be said of moral deference. As understanding-pessimists point out, many paradigmatic examples of moral deference are most easily interpreted as wholesale failures to pursue *either* of the aims of moral inquiry: when a person forms the belief that *p* but does not seek to understand why *p*, she neither pursues understanding of truth nor avoids apparent understanding of falsehood.[[16]](#footnote-16) Abandonment of a moral belief on the grounds that it might be false, by contrast, does not readily suggest such a shortcoming. Indeed, this behavior straightforwardly promotes one of the twin goals of moral belief: avoidance of merely apparent understanding of moral falsehoods.

Thus, even if understanding-pessimism successfully explains the problem with moral deference, it has only limited implications for moral disagreement.

1. Virtue

Robert Howell (2014) defends pessimism about moral deference by drawing on observations about moral virtue. This section surveys his brand of pessimism and explains why, like understanding-pessimism, it has only limited implications for cases of moral disagreement.

4.1 Howell’s Arguments for Pessimism

On Howell’s account, moral virtues crucially involve “dispositions to act, feel, and believe in certain ways” (2014, 403). A person who believes that she should act generously, but is not disposed to act generously, is not herself generous. The dispositions of the virtuous person, moreover, must be “unified” and “subjectively integrated” (403, 411).

Howell argues that this characterization of virtue illuminates several problems with moral deference. For one, moral deference can be a sign that a person already lacks virtue. If a person needs to defer about whether suffering is bad, Howell suggests, he “must be lacking intuitions, feelings, and other non-cognitive attitudes about suffering” (404). This can explain our discomfort with many of the examples of moral deference that are discussed by pessimists.[[17]](#footnote-17) Several such examples involve thinkers who begin their moral reasoning with a radically impoverished moral perspective—a perspective, Howell suggests, that is incompatible with virtue. Deference, then, can *reveal* a lack of virtue.

Howell places far more weight on a second line of argument: moral deference, as an alternative to reflective inquiry, can *interfere with* the development of virtue. He develops this theme through four related points. First, echoing understanding-pessimists like Hills, Howell argues that we cannot act virtuously on the basis of moral beliefs sustained through deference. The agent with “complete virtue” acts not merely in accordance with moral truths, but on the basis ofa reflective understanding of those truths (406). To the extent that an agent’s moral thinking is based on deference, it may well fail to provide her with such reflective understanding.

Second, Howell argues that deferential belief prevents the believer from achieving a virtuous reliability in thought and action. This is another idea prominent in Hills’s writing; both Howell and Hills worry that a deferential agent will not reliably “apply the belief to new, slightly different cases” in both thought and action (408).

Third, deference tends to introduce “cognitive isolation” (407). Beliefs sustained through deference are “isolated” in the sense that they are not based on the network of reasons, intuitions, and insights that grounds non-deferential moral beliefs. Rather, they are based on a sociological fact about other thinkers. Howell cites two problems with this contrast. It makes the agent less likely to detect “subtle forms of incoherence” across her deferential and non-deferential beliefs; it also makes her likely to act in a way that does not “reflect a coherent moral view” (407).

Finally, even when deference provides us with true beliefs, it tends not to provide us with the motivations, feelings, and intuitions required for moral virtue (405). An agent who can reach a true moral belief through reflection will be better-placed to attain virtue in that way than by gaining the relevant belief through deference.

It’s worth emphasizing the modesty of Howell’s pessimism. He does not argue that any of these problems is associated with every case of deference. And he also does not argue that these problems, together, always outweigh the benefits of deference. He even grants that, “[i]f one is in a position to learn a moral fact by deferring, and one cannot come to know that fact non-deferentially without substantial cost, it might well be that almost always one should do so” (390). Nevertheless, he insists, “there is something wrong with moral deference that is not in general wrong with non-moral deference” (392). While there are several ways in which deferential moral beliefs fall short of non-deferential ones, he argues, a deferential belief about the location of medium-sized object need not fall short of a non-deferential one in any sense.

4.2 Virtue and Disagreement

Suppose that, for the reasons Howell cites, moral deference really is in tension with a life of virtue. Does the same tension arise when a person revises her moral beliefs solely in response to disagreement? Well, at least one sort of response is guilty by association: as discussed in section 1, *switching beliefs* in response to disagreement simply is moral deference. To the extent that deference undermines virtue, this sort of response to disagreement does so as well.

I’ll now argue that Howell’s worries do not extend beyond this extreme response to disagreement. When a person responds to moral disagreement in a more moderate way—by suspending judgment or by merely reducing confidence in her beliefs—Howell’s concerns are toothless.

Let’s start with Howell’s first worry: that deference indicates an antecedent lack of virtue. This first point does seem apt to explain our discomfort with several canonical thought-experiments involving moral deference. But Howell does not place much emphasis on this point, and rightly so. Many cases of moral deference do not involve a thinker who starts from a troublingly impoverished moral perspective. And pessimists, Howell included, argue that deference is problematic even in cases that do not involve initial insensitivity to an obvious moral truth. So Howell’s first point is best seen as one of a suite of problems associated with moral deference; though it can explain our discomfort with certain cases, it does not by itself vindicate a robust pessimism about deference. Likewise, this point does not vindicate steadfastness about moral disagreement; though some moral disagreements involve troublingly impoverished moral views, many do not.

Howell places more weight on his second line of thought: he argues that, as an alternative to reflective inquiry, deference interferes with the development of virtue. This line of thought also fails to justify steadfastness about moral disagreement. The reasons to think that deference interferes with the development of virtue are not reasons to suspect that responsiveness to disagreement interferes with either the development of virtue or the abandonment of vice.

Begin with Howell’s first two ways of developing this theme, both of which echo understanding-pessimism. He argues that a policy of moral deference leaves an agent without the ability to act and believe for the right reasons, and that it also leaves her without the ability to apply her newly acquired belief to related cases. These are closely related points. They suggest that the virtuous agent’s outright beliefs come attended with a suite of abilities—the agent can, loosely speaking, put her outright beliefs to work in thought and action. But even if these claims about the abilities that should come along with belief are right, they do not imply anything about the abilities (or inabilities) that should come along with virtuous suspense of judgment. Nor do these claims call into question mere reductions in confidence; whether I am totally certain that *p* or only somewhat confident that *p*, I can act on the basis of my belief that *p*, and I can reason from my belief that *p* to appropriate conclusions about related cases. Policies of suspending judgment or reducing confidence in response to disagreement, then, present no threat to the virtuous agent’s ability to put her outright moral beliefs to work in thought and action.

Move on, now, to Howell’s thought that deference creates cognitive isolation. Howell cites two problems for the agent with cognitively isolated moral beliefs: it may be more difficult for her to notice subtle inconsistences between those beliefs, and since her actions may be directed by multiple incoherent approaches to ethics, they may be unpredictable and apparently random. He cites, in other words, two issues that arise when agents are partially committed to multiple conflicting approaches to ethics. But neither reducing one’s confidence nor suspending judgment in response to disagreement tends to lead agents into this worrisome position. Far from introducing new and potentially problematic commitments to an agent’s moral outlook, these responses to disagreement simply weaken extant commitments. They are far more likely to *resolve* subtle incoherence than to create it.

Finally, consider Howell’s point that moral deference is less likely to provide the believer with the motivations, intuitions, and emotions required for virtuous moral commitment. This point seems, at first blush, to provide the strongest case for a virtue-based steadfastness about moral disagreement. Perhaps, just as we are better-placed to acquire all the components of virtue through *a priori* reflection than through deference, we are better-placed to abandon all the components of vice—including vicious motivations, intuitions, and emotions—through *a priori* reflection than through responsiveness to disagreement.

But the comparison between these two lines of argument is not as apt as it may seem at first. To see this, note an important difference between the agent who defers and the agent who revises her views in response to disagreement. It’s usually a live possibility for a morally deferential agent that she is failing to perform a pattern of *a priori* reflection (albeit perhaps a very obscure and difficult one) that could, in principle, provide an alternate basis for her deferential moral belief. By contrast, it’s often *not* a live possibility for the agent who reduces confidence in response to disagreement that any particular pattern of *a priori* reflection could bring her to precisely the degree of belief revision that disagreement makes appropriate. Consider: it would be an odd, quixotic agent who set out to find arguments that would justify her becoming 20% less confident in a moral proposition rather than basing that 20% reduction on disagreement alone. Why think that any arguments exist that could successfully fill that role? By contrast, there’s nothing odd about an agent who has good evidence (say, from testimony) that a moral proposition is likely to be true, and then sets out to find the arguments that might generally justify one in believing that proposition.

This crucial contrast explains why Howell’s emphasis on motivations, emotions, and intuitions does not justify a general steadfastness about disagreement. Deferential belief may typically fall short of ideal, virtuous belief-formation; it usually fails to provide the emotions, motivations, and intuitions that would be provided by an alternate method of belief-formation, and that alternate method is a live alternative for the agent in question. But there is no way to levy a similar accusation at more modest responses to disagreement, precisely because there is quite often no salient alternative method of belief-revision. The precise shortcoming of moral deference, when it comes to cultivating virtuous dispositions, simply has no analogue in the case of responsiveness to disagreement.

We’ve now seen that none of Howell’s worries about moral deference justify analogous worries about non-deferential responses to disagreement; for all Howell says, responsiveness to disagreement might be a crucial part of the pursuit of virtue. Now, of course, abandoning beliefs in response to disagreement hinders the pursuit of virtue in at least one sense: when I abandon my true moral belief that *p*, I cannot use that belief in further reasoning. Nor can that belief be integrated into the network of beliefs and dispositions that is characteristic of virtue. But this does not show that the pursuit of virtue rules out suspense of judgment in response to disagreement. We saw in section 3.3 that any plausible account of the aim of moral belief must allow that, in order to avoid believing moral falsehoods, we can sometimes appropriately suspend judgment about some moral propositions. A parallel point holds when it comes to the pursuit of virtue. When we get adequately compelling evidence that some integrated suite of beliefs, attitudes, and motivations is not a virtue but a vice, we are morally beholden to abandon those beliefs, attitudes, and motivations. The importance of achieving virtue, like the importance of achieving understanding, cannot explain why it would be worse to do so on the basis of disagreement than on the basis of private reflection.

To conclude: Howell’s line of thought may successfully explain why moral deference is distinctively problematic. Perhaps the best way to strive toward the network of dispositions crucial to a life of virtue is to grow and sustain those dispositions together, and not in the piecemeal way that deference seems to involve.[[18]](#footnote-18) By contrast, there is no reason to think that the best way to pursue the crucial goal of avoiding vice involves giving no weight to others’ beliefs.

1. Conclusion: A Final Objection

Intuitively, there is a straightforward connection between pessimism about moral deference and steadfastness about moral disagreement. If we should not form or sustain moral beliefs on the basis of what other people think, it seems reasonable that we should also give their moral thinking little to no weight when they disagree with us. This paper has shown that this line of thought is too hasty. The most attractive, well-developed forms of pessimism about moral deference are compatible with a quite optimistic approach to responsiveness to disagreement.

But just how attractive would it be to combine these two positions? Suppose, as Hills suggests, that the problem with moral deference is an epistemic one. Suppose, further, that there is no principled picture on which deference is epistemically forbidden but other forms of responsiveness to disagreement are epistemically permissible. This would provide a serious challenge, based in considerations of theory-selection, to the compatibility thesis I have defended. So, are there any non-*ad hoc* pictures of moral epistemology on which belief-revision in responsiveness to others’ views could be appropriate up to the point of suspending judgment, but no further?[[19]](#footnote-19)

I think there are likely multiple pictures on which this policy could be explained in a principled way. I’ll close by briefly sketching one such picture, which relies on the gap between degrees of belief (or credences) and outright belief. Suppose that the epistemic norms that govern response to disagreement are norms on credences; when faced with disagreement, one should alter one’s credences precisely to the extent that one’s interlocutor has certain epistemic credentials (for instance, a good track-record in the relevant domain of inquiry). Suppose further that credences and outright beliefs are importantly distinct, such that the mere fact that I have a certain credence between 0 and 1 often does not suffice to determine whether I have an outright belief. On this picture, the epistemic norms governing whether to form an outright belief or to suspend judgment would not be reducible to the epistemic norms governing credences. In this way, the fact that disagreement demands a certain degree of credence revision might, in many cases, fail to determine whether that revision demands the formation of an outright belief. On this picture, even if moral disagreement frequently requires quite a bit of credence revision, the norms on formation of outright belief might be (for one reason or another) extremely stringent when it comes to moral matters. This could explain why we are seldom in a position to form an outright moral belief based only on disagreement.[[20]](#footnote-20) This schema for epistemology, then, can accommodate pessimism about moral deference while resisting, in a principled way, steadfastness about moral disagreement.

It would be too quick, then, to suppose that considerations of theory-selection push all pessimists about moral deference to embrace steadfastness about moral disagreement. The most attractive, well-developed versions of pessimism about moral deference leave open the possibility that taking others’ moral views into account can, in at least some important forms, be both morally and epistemically appropriate.

**References**

[Christensen](http://philpapers.org/s/David%20Christensen), David. 2009. “Disagreement as Evidence.” *Philosophy Compass* 4 (5): 756-767.

---. 2011. “Disagreement, Question-Begging, and Epistemic Self-Criticism” *Philosopher’s Imprint* 11 (6): 1-22.

Decker, Jason and Daniel Groll. 2014. “Moral Testimony: One of These Things Is Just Like the Others.” *Analytic Philosophy* 54 (4): 54-74.

[Elga](http://philpapers.org/s/Adam%20Elga), Adam. 2007. “Reflection and Disagreement.” *Noûs* 41 (3): 478–502.

Enoch, David. 2014. “A Defense of Moral Deference.” *Journal of Philosophy* 111 (5): 229-258.

Hazlett, Allan. 2015. “The Social Value of Non-Deferential Belief.” *Australasian Journal of Philosophy* 94 (1): 131-151.

Hills, Alison. 2009. “Moral Testimony and Moral Epistemology.” *Ethics* 120 (1): 94-127.

---. 2010. *The Beloved Self: Morality and the Challenge from Egoism*. Oxford: Oxford University Press.

---. 2013. “Moral Testimony.” *Philosophy Compass* 8 (6): 552-559.

---. 2015. “Cognitivism about Moral Judgment.” In *Oxford Studies in Metaethics,* vol. 10, edited by Russ Shafer-Landau, 1-25. Oxford: Oxford University Press.

Hopkins, Robert. 2007. “What is Wrong with Moral Testimony?” *Philosophy and Phenomenological Research* 74 (3): 611-634.

Howell, Robert J. 2014. “Google Morals, Virtue, and the Asymmetry of Deference.” *Noûs* 48 (3): 389-415.

Jones, Karen. 1999. “Second-Hand Moral Knowledge.” *Journal of Philosophy* 96 (2): 55-78.

Lackey, Jennifer. 2008. “What Should We Do When We Disagree?” In *Oxford Studies in Epistemology*, vol. 3, edited by Tamar Gendler & John Hawthorne, 274-93. Oxford: Oxford University Press.

---. 2010. “A Justificationist View of Disagreement’s Epistemic Significance.” In *Disagreement*, edited by R. Feldman and T. Warfield, 298-325. Oxford: Oxford University Press.

McGrath, Sarah. 2009. “The Puzzle of Pure Moral Deference.” *Philosophical Perspectives* 23 (1): 321-344.

---. 2011. “Skepticism about Moral Expertise as a Puzzle for Moral Realism.” *Journal of Philosophy* 108 (3): 111-137.

McConnell, Terrance C. 1984. “Objectivity and Moral Expertise.” *Canadian Journal of Philosophy* 14, (2): 193-216.

Mogensen, Andreas L. 2015. “Moral Testimony Pessimism and the Uncertain Value of Authenticity.” *Philosophy and Phenomenological Research* 92 (3): 1-24.

Nickel, Philip. 2001. “Moral Testimony and its Authority.” *Ethical Theory and Moral Practice* 4 (3): 253-266.

Pasnau, Robert. 2015. “Disagreement and the Value of Self-Trust.” *Philosophical Studies* 172 (9): 2315-2339.

1. McGrath (2011, 115), Mogensen (2015, 5), and Howell (2014, 394) argue for framing pessimism as a thesis about *deference* rather than *testimony*. [↑](#footnote-ref-1)
2. For the purposes of this paper, I set aside non-cognitivist views of moral judgment and refer to moral judgments as ‘beliefs.’ See Hopkins (2007, 617-8) for reasons to doubt that non-cognitivism offers a distinctive, plausible vindication of pessimism. [↑](#footnote-ref-2)
3. Howell (2014, 390) argues for the possibility of deference to one’s own past or future beliefs. I set this complication aside. [↑](#footnote-ref-3)
4. See, for example, Lackey (2008, 2010), Christensen (2011, 8-9) and Elga (2007, 488-90). [↑](#footnote-ref-4)
5. This line of argument is developed further in McGrath (2011, 118-20) and Decker & Groll (2014, 71-2). [↑](#footnote-ref-5)
6. A nearby strategy stresses *expertise*; see McGrath (2011, 126-30). [↑](#footnote-ref-6)
7. Conciliationists differ as to what it means to be “reliable.” Elga (2007, 490) emphasizes accuracy; Christensen (2011, 15) asks whether others are well-informed and likely to have reasoned well. My discussion is neutral between these options. [↑](#footnote-ref-7)
8. See Hopkins (2007, 620-1) and McGrath (2011, 129-30). [↑](#footnote-ref-8)
9. McGrath (2009, 323) is likely an exception. But note that McGrath (2011, 129) offers distinct reasons to look beyond the reliability strategy. [↑](#footnote-ref-9)
10. Mogensen expresses doubts about the “true self” and the value of authenticity (2014, 15-6). [↑](#footnote-ref-10)
11. Jones (1999, 57) quotes Wolff only to refute him; Mogensen (2015) expresses skepticism toward both value of authenticity and pessimism about moral deference. [↑](#footnote-ref-11)
12. According to Hills, even deference regarding a proposition of the form *q is why p* does not put a person in a position to grasp the reasons why *p*, and therefore does not suffice to provide understanding (2009, 101). [↑](#footnote-ref-12)
13. Cf. Christensen (2009, 763). [↑](#footnote-ref-13)
14. Though Hills (2010) argues against suspending judgment in response to disagreement, she does not take a clear position on mere reductions in confidence. [↑](#footnote-ref-14)
15. By ‘apparent understanding,’ I mean the state that is exactly like understanding except that it is not factive. [↑](#footnote-ref-15)
16. Hazlett (2015) raises an important objection to this claim: couldn’t a seeker of moral understanding frequently begin her search through deference, and then seek to turn deferential beliefs into understanding? [↑](#footnote-ref-16)
17. See, for instance, Hills’s case of the Knowledgeable Extremist (2009, 115). [↑](#footnote-ref-17)
18. But see Mogensen (2015, 12-15) for objections. [↑](#footnote-ref-18)
19. Thanks to Declan Smithies for this objection. [↑](#footnote-ref-19)
20. A caveat: on most models of this sort, disagreements requiring credence 1.0 in an opponent’s beliefwill also require outright belief that *p*. They will, then, require deference. But the most plausible forms of pessimism allow that extreme cases (as when a child defers to a parent) make moral deference appropriate. Cases that rationalize credence 1.0 in an opponent’s view are liable to seem extreme in just this way. [↑](#footnote-ref-20)