PHILOSOPHICAL TOPICS
VOL. 38, NO. 2, FALL 2010

Autonomy, Understanding, and Moral Disagreement

C. Thi Nguyen
Utah Valley University

ABSTRACT: Should the existence of moral disagreement reduce one's confidence in one's moral judgments? Many have claimed that it should not. They claim that we should be morally self-sufficient: that one's moral judgment and moral confidence ought to be determined entirely one's own reasoning. Others' moral beliefs ought not impact one's own in any way. I claim that moral self-sufficiency is wrong. Moral self-sufficiency ignores the degree to which moral judgment is a fallible cognitive process like all the rest. In this paper, I take up two possible routes to moral self-sufficiency.

First, I consider Robert Paul Wolff's argument that an autonomous being is required to act from his own reasoning. Does Wolff's argument yield moral self-sufficiency? Wolff's argument does forbid unthinking obedience. But it does not forbid guidance: the use of moral testimony to glean evidence about nonmoral states of affairs. An agent can use the existence of agreement or disagreement as evidence concerning the reliability of their own cognitive abilities, which is entirely nonmoral information. Corroboration and discorroboration yields nonmoral evidence, and no reasonable theory of autonomy can forbid the use of nonmoral evidence. In fact, by using others to check on my own cognitive functionality, an agent is reasoning better and is thereby more autonomous.

Second, I consider Philip Nickel's requirement that moral judgment proceed from personal understanding. I argue that the requirement of understanding does forbid unthinking obedience, but not discorroboration. When an agent reasons morally, and then reduces confidence in their judgments through discorroboration, they are in full contact with the moral reasons, and with the epistemic reasons. Discorroboration yields more understanding, not less.

What are we to do about moral disagreement? A great many people disagree with me about basic moral issues, and I know that many of them are very thoughtful, reflective, sensitive people. In empirical matters, we treat such disagreements as important.[1] If I disagree with my girlfriend about whether we left the oven on, this gives both of us a reason to doubt our memory. But many of us think that moral disagreement is different; that moral disagreement is something we should ignore, unlike empirical disagreement.

One common view is that moral disagreement in and of itself doesn't matter. Let me call this view *moral self-sufficiency*. This is the view that our moral beliefs ought to follow from reasoning we ourselves have done, by reasons we see, understand, and accept. Other people may point out new arguments and bring up new information for us to consider, but our moral beliefs should eventually be determined entirely through our own reasoning. In contrast, we do not take ourselves to be self-sufficient in empirical domains; we are free to trust doctors, mechanics, and historians without doing the reasoning for ourselves. Moral self-sufficiency is the view that, though empirical cognition depends on trusting others, moral judgment should be ultimately independent.

I think that moral self-sufficiency is wrong. Moral self-sufficiency ignores the degree to which moral reasoning is a fallible cognitive process like all the rest. I argue instead for what I call *moral humility.* I claim that the existence of certain types of moral disagreement should bring us to doubt our moral judgments, to some degree. I'm going to talk here only about the impact of moral disagreement, but a larger question looms in the background. Is moral cognition subject to the same epistemic regulations of conduct as other cognitive domains? Or does moral life get a special exemption from the normal rules of proper epistemic conduct?

In this paper, I will consider several distinctively moral arguments for moral self-sufficiency.[2] These arguments will all claim that moral judgment gets a special exemption from the usual methodology of discorroboration, for distinctively moral reasons—that the use of moral testimony is morally wrong. This is not to say that moral testimony is bad evidence. Rather, moral testimony might be perfectly good evidence, but it is evidence that is wrong to use. This thought should be familiar; it's like improperly gathered police evidence at a criminal trial. A bloody knife, taken from your home without a search warrant, may be good evidence of your guilt—but it's evidence we ought to ignore.

I will take up two such arguments for moral self-sufficiency. First, I will consider arguments from Robert Paul Wolff that any use of moral testimony is a vio-

lation of the constraints of *autonomy*. Second, I will consider Philip Nickel's claim that the overuse of moral testimony violates a special requirement for *understanding* in the moral domain. I hope to show that, though each of these worries captures an important consideration for moral life, none of them is so powerful as to block the effect of moral disagreement. As long as we take ourselves to be cognitively fallible beings, and as long as part of our moral enterprise is to *get it right*, then we must be open to new evidence of errors in our thinking.

I. WOLFF'S ACCOUNT OF AUTONOMY

First, I'd like to focus here on the version of this argument that seems to me the most intuitively plausible—the argument from autonomy.[3] The argument, loosely, is that since we are free, autonomous beings, we ought to decide for ourselves on moral matters. Any significant use of moral testimony is therefore a violation of our autonomous nature. There are many different versions of this argument, using subtly different conceptions of autonomy. Unfortunately, I haven't come close to settling my mind about which conception is right. But I suspect that any reasonable conception of autonomy will be compatible with my epistemic arguments. The best I can do is to take up the most radical conception of autonomy, and show that, even there, disagreement must count.

Robert Paul Wolff presents, in *In Defense of Anarchism*, one of the strictest autonomy requirements out there. In fact, he thinks that the requirements of autonomy requires us to be anarchists. But I think, even if I grant Wolff's very strident view of autonomy, I can show that moral disagreement still matters.

Let me begin by defining some terms. I can be said to be "using testimony" when an agent's testifying that p gives me some reason to believe p. "Moral deference," in contemporary usage, is a particular use of testimony.[4] Sarah McGrath defines deference as a case in which "one holds a view solely because another person holds that view."[5] Problems surrounding moral testimony have also been examined as issues in *moral authority*. Elizabeth Anscombe defines "moral authority" as "taking that somebody said something as over and above what you decide for yourself."[6]

Wolff argues that we should ignore all moral testimony, and so ignore disagreement, on considerations of autonomy. A genuinely autonomous, responsible moral agent must reason for himself or herself and act from his or her own judgment. To act otherwise, thinks Wolff, is to make yourself a slave—to act merely on another's say-so.

Wolff's requirement for autonomy takes the form of two requirements. First, accepting the *command* of another, for example, constitutes a violation of autonomy, says Wolff:

> Since the responsible man arrives at moral decisions which he expresses to himself in the form of imperatives, we may say that he gives laws to

> himself, or is self-legislating. In short, he is *autonomous*. As Kant argued, moral autonomy is a combination of freedom and responsibility; it is a submission to laws which one has made for oneself. The autonomous man, insofar as he is autonomous, is not subject to the will of another. He may do what another tells him, but not *because* he has been told to do it . . . . Inasmuch as moral autonomy is simply the condition of taking full responsibility for one's actions, it follows that men can forfeit their autonomy at will. That is to say, a man can decide to obey the commands of another without making any attempt to determine for himself whether what is commanded is good or wise.[7]

We may not act on the command of another because we have been commanded; rather, we must determine for ourselves whether our action is good. I'll call this the requirement for *self-legislation*.

Wolff provides a second requirement, directed at the use of moral testimony:

> The responsible man is not capricious or anarchic, for he does acknowledge himself bound by moral constraints. But he insists that he alone is the judge of those constraints. He may listen to the advice of others, but he makes it his own by determining for himself whether it is good advice. He may learn from others about his moral obligations, but only in the sense that a mathematician learns from other mathematicians—namely by hearing from them arguments whose validity he recognizes even though he did not think of them himself. He does not learn in the sense that one learns from an explorer, by accepting as true his accounts of things one cannot see for oneself.[8]

I'll call this the requirement for *moral understanding*. This is the requirement that when we self-legislate, we do so from reasons and reasoning that we understand and accept.

These two requirements for self-legislation and moral understanding together I will call Wolff's requirement of autonomy. It may seem as if the autonomy requirement will bring us to ignore moral testimony in all forms. Let me grant, for the moment, that there is such an autonomy requirement. I'm going to argue that, even if Wolff is right about the autonomy requirement, it doesn't actually forbid all uses of testimony. I will argue that the autonomy requirement does forbid outright obedience to testimony, but it doesn't forbid a subtler use of testimony—the use of testimony in corroboration and discorroboration through agreement and disagreement. In fact, I think these corroborative uses of testimony actually increase autonomy.

1.1 OBEDIENCE AND GUIDANCE

Let's start by clarifying the distinctions between certain healthy and unhealthy relationships between autonomous agents and testimony. First, autonomy considerations can only forbid very particular relationships to testimony. In Wolff's Kantian language, autonomy considerations forbid giving oneself a law or rule based solely on the testimony of another. Autonomy considerations cannot forbid taking up

normatively neutral information through testimony. It can't be a crime against autonomy to ask another person for directions.

What it is to be rational is to believe in accord with the reasons and evidence; in that important sense, we rational beings are not free to believe as we wish. The mental freedom for autonomous beings is freedom in a very specific domain; it is freedom about which practical rules for action we give ourselves. It is not the freedom to believe as we please. After all, if a contemporary scientist believes that 2 + 2 = 5, or that the Earth is flat, the scientist isn't free—he's dumb. Whatever autonomy is, it must be compatible with these conditions of rationality. Taking evidence is not the same as taking orders.

What exactly is the difference, then, between the healthy way in which an agent's belief can be determined by evidence, and the unhealthiness of excess obedience? Let's begin with Wolff's paradigmatic case of bad obedience, the unthinking soldier. We are to imagine a soldier who obeys wartime orders without deciding for himself if they are morally permissible. The fact that he is ordered by his superior officer is sufficient to justify action. I will assume that the unthinking soldier does represent a genuine problem for autonomy.

But what exactly is wrong with unthinking obedience? It's actually rather tricky, for there are superficially similar cases that aren't problematic.[9] Take, for instance, the case of following somebody in a car. Imagine that my friends and I are coming back from a camping trip, in which we are divided into two cars. We wish to drive into town and meet up at a restaurant. Our cell phones have run out of power, and we are in unfamiliar territory. The sensible plan is for one car to be the leader, and the other the follower, and for the driver of the first car simply to decide, and the driver of the second car simply to follow. Surely this case is morally unproblematic, but what, precisely, is the difference between the driving case and the unthinking obedience case?

Some have suggested that bad obedience is characterized by a lack of understanding in the initial decision to follow. But this can't be it, because an unthinking soldier may understand why he decided to follow. ("He is my superior officer, I'm in the military.") Perhaps, then, bad obedience is distinguished by the failure to understand the particular reasons for the subsequent particular actions and decisions. But that can't be it, either, since, in the driving cases, the follower also doesn't understand why he's turning right or left in each particular instance.

Perhaps the difference lies in the relationship of the follower to the overall goal. While the car-follower doesn't understand the particular micro-rules and decisions being used for each turn, he does understand the particular end being pursued. Perhaps the problem with the unthinking soldier is that once he's decided to follow, he doesn't understand the overall goals of the actions he's been ordered to do.

But I don't think that's quite it, either. Imagine that I come home on my birthday and my girlfriend blindfolds me and tells me to follow her. I ask why, and she says, "It's a surprise." She leads me out of the house, and I follow blindly. In this case, I am following the direction of another without a clear understanding of and

assent to the overall goal. I merely trust that my girlfriend has my best interests at heart. But this case is also obviously morally unproblematic.

Might it be, then, that unproblematic cases involve only limited obedience, and the unthinking soldier has unlimited obedience? Presumably, when we follow a lead driver, we won't follow them *anywhere*—not over the side of a bridge, for example. But this can't be it, either, for the unthinking soldier need not be obedient in every domain. It is safe to imagine that an unthinking soldier wouldn't do absolutely *anything*—most probably would not follow orders to shoot themselves in the head, or shoot their commanding officer from behind, or to burn down their grandparents' retirement home. The unthinking soldiers may be unthinking only in a particular domain—about their actions in a foreign country or toward enemy soldiers or a particular ethnic group—and yet still be problematically unthinking.

The problem lies in the nature of each individual instance of obedience. Obedience always involves an attempt to surrender responsibility for an action. Suppose I'm following my friend's car, and he runs a red light. If I run the light, too, I cannot excuse myself from responsibility by saying that I was merely following. I am responsible for that decision. The difference is that in the car-following case I am being *guided*, and in the unthinking soldier case the soldier is *obeying*. In guidance, the leader is simply *suggesting* a course of action, to which the follower assents or refuses. The assent may be invisible, because it is often so quick and such a simple judgment. In car-following, the actions—turning left, changing lanes—are all so utterly routine that the moral endorsement is virtually unnoticeable. But we can see that there *is* assent in each case, simply by imagining what happens when the leader car does something a little dicey—for example, running a red light. The following driver must decide whether or not he will follow, and he is responsible for his decision. In the problematic cases of obedience, I act on a rule solely because it is endorsed by another, or endorse a rule on the sole grounds that another endorsed it. In the unproblematic cases of guidance I uptake only *information*, and provide or withhold endorsement on my own recognizance.

The distinction between obedience and guidance captures something very important about where genuine autonomy problems lie, and where they don't. In the unproblematic cases, I am merely being guided: I am accepting only information from the leader, and then deciding for myself. In the problematic cases, I am obeying: I am accepting rules from the leader, which include information, and decisions about appropriate ends to pursue and methods to pursue them. I am letting them decide for me. Guidance may resemble obedience closely, but they are crucially distinct. I may modify my beliefs and actions in response to testimony, but this in and of itself does not show problematic obedience. What matters is whether my own judgment and endorsement is interposed between the testimony and the action. When another provides only information, and I provide the rules and judgment that connect that information to action, then any following I'm doing is unproblematic, though it may superficially resemble obedience.

What I've argued so far is that autonomy restrictions cannot block the use of strictly nonmoral information. This argument depends heavily on a strong distinction between moral and nonmoral content. Now, I don't take myself to have shown exactly what the line is between moral and nonmoral content. There are plenty of border cases that I cannot resolve. But I think there are some clear cases, and they're all I need to make my case.

So: there are at least two different ways we can interact with a given piece of moral testimony. We can obey the testimony, and so violate the autonomy requirement. Or, we can also use that testimony to glean nonmoral information. This is the use of moral testimony for corroborating and discorroborating our cognitive abilities using agreement and disagreement with others—for self-checking. This use of testimony, I claim, is obviously nonmoral, and so unproblematic.

I can use testimony to establish that there is disagreement, and the existence of disagreement can give me a reason to suspect the reliability of my own abilities. This piece of reasoning is entirely nonmoral. Evidence that my faculties are possibly unreliable is not a rule for action; it is not a moral fact of any sort. It is information about a nonmoral matter: the proper functioning of my cognitive faculties. Thus, even if we subscribe to the strongest prohibition against uptaking any form of endorsement or moral rule, there is still a pathway for moral disagreement to permissibly affect our confidence.

This claim—that there is a nonmoral use of moral testimony—is the most important claim in my argument. Let me slow down here and give my reasons in some detail. When I encounter a piece of moral testimony, there are at least two ways I can interact with it. I can simply obey. I can also use testimony as psychological information—merely to establish what a testifier believes. The normative use may be forbidden, but the informational use is clearly nonmoral and thus permissible. There are, accordingly, two ways in which testimony can be used to modify my own belief; I will call them the direct pathway and the indirect pathway. In the direct pathway, I come to alter my belief immediately in accordance with received testimony. This is obedience. In the indirect pathway, I use the testimony to establish what my interlocutor believes, then use the fact that we agree or disagree as evidence in evaluating our reliability.

We don't often separate these two uses of testimony in everyday empirical matters, since both uses of empirical testimony are permissible, but the distinction is crucial in the case of moral testimony. In the direct pathway, I come to believe a moral rule solely on the basis of another's believing it. In the indirect pathway, I am using testimony to glean a piece of morally neutral information: that an interlocutor I take to be reliable on the subject agrees or disagrees with me. The distinction between the two pathways is everyday. When I'm doing my math homework, I am working under the requirement that my answers and my work are my own. It does violate that requirement to simply copy my partner's work and conclusion. But it

surely doesn't violate that principle to check my work against other students, and to doubt my work (and theirs) when there is disagreement.

To understand the difference between the two pathways, it is crucial to distinguish between two processes of reasoning; I'll call them the moral reasoning process and the self-checking process. The moral reasoning process begins from considerations about what's good and bad, and concludes with a comment about how I should act. The self-checking process begins from considerations about my coherency, reliability, and evidence of error, corroboration, and discorroboration, and ends with a conclusion about how I ought to trust the moral-reasoning process. It starts from nonmoral information, yields a nonmoral conclusion, and proceeds by reasoning I understand and accept. Furthermore, when I act on that information—when I change my own beliefs—I do so according to other rules about justification and belief revision that I understand and endorse: namely, that disagreement between reliable agents is a reason to doubt both.

Obedience—the direct pathway—is an attempt to skip over the moral reasoning process. But the indirect pathway does not run afoul of the autonomy requirement, because it involves engaging in the moral reasoning process and the self-checking process. When I corroborate and discorroborate, I don't skip over a moral reasoning process. I perform the entire moral reasoning process, and then perform another elaborate self-checking process on top of it.

So self-checking isn't moral reasoning; thus it won't run afoul of the autonomy requirement. The self-checking process will, however, impact my confidence in the results of the moral reasoning, and so result in changes to my all-things-considered beliefs. This may seem paradoxical, but it is not. This conclusion depends on the claim that there are empirical facts that lie under and buttress my confidence in my moral reasoning; thus, nonmoral reasoning can infect my confidence in my moral reasoning. But surely this claim is correct. Empirical evidence can impugn my trust in my own judgments. Surely, if I were to find out that I had been under the influence of drugs, or post-traumatic stress syndrome, when I passed moral judgment on my friends, I would have reason to worry about and doubt those judgments. Since in all reasoning, including moral reasoning, we are putting our trust in defeasible reasoning processes and abilities, it seems difficult to see why there *couldn't* be empirical claims that could impugn that trust.

This may seem like hair-splitting. It is not. I am trying to show that the process of obedience and the process of discorroboration are entirely different; that what we find abhorrent about blind obedience is nowhere to be found in discorroboration. When we obey unthinkingly, we take ourselves out of the reasoning process. We cannot give any explanations for the rules we follow; we are not involved in the process of weighing and deliberating and deciding. We make at most a single decision—to obey an authority—and then step out of the picture. When we corroborate and discorroborate, we do the opposite. When we check our reasoning against others, we are buried in the reasoning process. We are fully engaged in two distinct forms of reasoning—first, forming our mind independently about the moral mat-

ter, and then checking the reliability of that first process through another process. Both processes proceed in the full light of understanding. I know why I've decided that eating meat is allowed (because I think animal lives are, in the end, genuinely less important than human desires), and I know why I hesitate to act on that belief (because not everybody agrees with me, and because this is a reason to worry). I am in full contact with the reasons and rules involved; I am fully a *participant*, and not a resigned bystander, in the reasoning process.

If the process of corroboration and discorroboration is not essentially a moral matter, why does it trip alarms in the vicinity of autonomy? Let me offer a diagnosis. The autonomy-based worries about disagreement depend on drawing a false dichotomy between our moral commitments and our epistemic commitments. The autonomy worries arise from seeing the epistemic pressure from discorroboration as an alien and invasive presence in our moral lives. The view depends on seeing our relationship with our moral beliefs as more genuine and integral than considerations of cognitive reliability. When procedures like self-checking and discorroboration get in between my private moral experience and my all-things-considered beliefs, it may seem like I'm alienating myself from my moral commitments. But self-checking proceeds from other commitments I myself have made—epistemic commitments, my commitment to getting it right. As long as our moral beliefs claim to objectivity, the friction here is one that's entirely internal to moral life.

I can put things a little more clearly if I help myself to the Kantian framework which seems to lie underneath Wolffian autonomy worries. Let me suppose, for the moment, that there is a very strong connection between rationality and autonomy. We are rational beings, and in reasoning well, we are autonomous. The unthinking soldier is not reasoning at all; he is submitting. But the epistemically motivated doubter is not giving up on thinking. She is reasoning more, she is taking up more responsibilities in the quest to reason well. Insofar as autonomy is, as Kant suggests, performing the right reasoning for yourself, by attending to disagreement we become *more* autonomous.

1.3 DIAGNOSIS: BIVALENCE AND ALIENATION

If the process of corroboration and discorroboration is not essentially a moral matter, why does it trip alarms in the vicinity of autonomy? Let me offer a diagnosis. The autonomy-based worries about disagreement depend on drawing a compelling, but false, dichotomy between our moral commitments and our epistemic commitments. The autonomy worries arise from seeing the sort of rational, epistemic pressure from discorroboration as an alien and invasive presence in our moral lives. The view depends on seeing our relationship with our moral beliefs as more genuine and integral than considerations of cognitive reliability. When procedures like self-checking and discorroboration get in between my private moral experience and my all-things-considered beliefs, it may seem like I'm alienating myself from my moral commitments. But this view suppresses the complexly cognitive

nature of moral beliefs, and our complex nature as rational autonomous beings. Our moral commitments are an essential part of our personal identity, but so are our epistemic commitments. We are cognitive beings as well as beings soaked in moral feeling. When we encounter disagreement and lose confidence in ourselves, we are not being alienated through an invasion from without; we are in conflict with another part of ourselves.

The autonomy worry about discorroboration arises from a one-dimensional view of moral beliefs: the view that our moral beliefs are simply products of our freedom, as expressive of our nature and phenomenal experience of the moral world. As long as we are moral cognitivists, we must view moral beliefs as essentially *bivalent*.[10] We are committed both to a moral belief's being expressive of ourselves *and* as getting it right of facts independent of ourselves. The first set of commitments drive us to keep our moral beliefs private, to seal them off from social input. But the commitment to correctness drives us in the opposite direction; it drives us to procedures of corroboration and discorroboration, to general epistemic principles that obtain of any cognitive domain. Neither of these commitments are alien; both sets of commitments arise from the nature of moral beliefs themselves. Insofar as we take our moral beliefs to be aimed at truth, we commit ourselves to using procedures oriented toward accuracy and reliability.

It does seem that there is friction between our commitments to moral expressiveness and our commitments to moral accuracy, but this friction is internal to our body of commitments. We wish to rule ourselves by rules we give ourselves, and we wish to give ourselves rules that are in line with our felt, phenomenally vivid moral experiences, but we also wish to give ourselves the right rules. Conflict is inevitable, so long as our moral feelings are simultaneously personal and purportedly cognitive.

The same friction can be seen elsewhere. I wish, for example, to be with a romantic partner that I feel love for. But I also wish to love properly. I have learned—through trial, error, and lots of advice from friends—that I tend to be attracted to sadistic narcissicists, and this attraction easily blooms into love. So I walk a difficult path—I try to stay true to my feelings, but also stand back from them, make sure they're right. I cannot inhabit my loves without conflict, for my loves are enmeshed in complex, sometimes contradictory commitments. This conflict is not an invasion from without, for it arises from the fact that there are pressures from different directions laid over my love. The friction is not a sign of alienation, but of the complexity and difficulty of human commitments.

1.4 FIRST-ORDER CONTENT AND SECOND-ORDER REFLECTION

I am certainly not arguing that all uses of testimony are legitimate. I grant that the direct pathway is an autonomy-violating use of testimony. And the direct path is the only way to get new moral content through testimony. The indirect path can only manipulate the second-order degree of confidence I have in that first-order moral judgment.

So this is how I think it goes. I think eating meat is all right, my trusted ethics advisor thinks it's wrong. I don't accept her argument, but since she's so smart and so thoughtful, I grant that her contrary belief gives me a moderate reason to suspect that I've reasoned wrongly. I don't acquire from her the belief that I ought to be a vegetarian. I acquire from her evidence that I could be unreliable. Thus I have reasoned to a moral belief—that eating meat is wrong—by my own reasoning, preserving my autonomy. And I have reasoned to the belief that that last bit of reasoning might have a mistake in it, also using my own reasoning, using only nonmoral evidence I acquired through testimony. And this second bit, I claim, is also autonomy-preserving.

Thus, I think, if the autonomy restrictions are right, we can acquire new moral content only through our own reasoning, but we can adjust our self-confidence in the acuity and accuracy of our reasoning process by using testimonial evidence. This will have implications for action, for the justification of our actions often depends not only on the content of our belief, but the degree of our confidence. Most of us think, for example, that a drastic, irrevocable action—such as killing another person—depends not only on believing that it's right, but on being very sure that it's right.[11]

This dual reflective stance shows up in other places; most of us hold such a position toward our judgments of romantic love. We seem to hold a principle of the autonomy of love. We think that when we love somebody, that love must be arrived at through a private process of deliberation and feeling. Nobody can tell me *who* to love; that is something that only I can come to discover for myself. But, at the same time, my process of finding love, falling in love, and accepting my loves is subject to input from other people. I know that I make mistakes in love. I've learned that I have a tendency to fall for sadistic narcissists, and I never seem to realize when I'm doing it again, though it's obvious to every single one of my friends. Though only I may come to love another, others can certainly help me figure out when I'm loving badly, or when I'm mistaken about my feelings, or under the influence of neurosis, or addled by drugs or lust. I can be reminded by others that my judgment is impaired, that I'm recently out of a relationship and my emotions are wild, or that I've had too much to drink that night and am clearly drunk and so shouldn't get married just yet. Only I can decide who I love, but others can suggest to me that my judgment is, at the moment, impaired.

1.5 RATIONAL SELF-PERFECTION

What I've been arguing is that any reasonable theory of autonomy *must allow a rational autonomous being to uptake information without thereby violating her autonomy*. Uptaking morally neutral information cannot count as an autonomy violation. When I use the existence of disagreement to reflect about my own reliability, I am a full participant in the reasoning process.

I can put things a little more clearly if I help myself to more of the Kantian framework which seems to lie underneath Wolffian autonomy worries. Let me

suppose, for the moment, that there is a very strong connection between rationality and autonomy. We are rational beings, and in reasoning well, we are autonomous. The connection is likely the animating force behind Wolff's worries. The unthinking soldier is not reasoning at all; he is submitting. He is giving up on his duties as a rational being to think, consider, and take evidence and reasons into consideration. But the epistemically motivated doubter is not giving up on thinking. She is reasoning more and insofar as she is reasoning from good epistemic principles, she is reasoning *better*.

Once we begin to take into account the moral beliefs of others—not by submitting to the command of others, but by becoming aware of the reasoning of others and using it as evidence—we become more rational. We are engaged in further rational activities: we are triangulating, we are corroborating, we are error-checking, we are debugging. We are engaged in an activity of increasing the reliability of our judgments. We are searching for evidence of our errors, evidence that may lead us to feel out where we should think again, where we should worry, where we should theorize. Not all instances of interaction with testimony are bad. When we use testimony to become more accurate, more well-informed, more self-conscious reasoners, we are increasing our autonomy. Other people can help us to become more autonomous, as long as we use evidence of their testimony thoughtfully and properly, and not lazily. The unthinking soldier has abandoned interaction with standards of correctness, while she who attends to disagreement is thereby pursuing correctness more fervently.

II. MORAL UNDERSTANDING AND THE RIGHT SORT OF REASONS

Another significant branch of the contemporary discussion seeks to explain the wrong of moral deference by invoking a duty for moral understanding. There is, according to these accounts, a duty to understand one's reasoning behind one's own moral beliefs and judgments. Moral deference and other excessive uses of moral testimony are wrong because they replace moral understanding with trust. When an agent undertakes actions solely on the say-so of another, they may get the action crudely right, but they do not do so with a full understanding of why the action is right.

There likely is an independent duty for moral understanding, and some of the contemporary accounts of the duty seem as plausible as the autonomy accounts.[12] But, I argue, any plausible version of the duty for moral understanding will permit the use of moral disagreement. In fact, I claim, the most plausible account of the duty will actually encourage us to attend to disagreement for the very same reasons that we are required to understand. Moral understanding and self-checking are both vital parts of the process of becoming the most accurate, reliable, morally sensitive beings we can.

2.1 HOPKINS AND GRASPING THE MORAL GROUNDS

Let us presume that there is something wrong with outright moral deference. Furthermore, let us grant that the *reason* moral deference is wrong is that it engenders a systematic lack of moral understanding. But does weighting disagreement also engender a lack of moral understanding?

Robert Hopkins offers an excellent overview of the problem of moral deference in his paper, "What Is Wrong with Moral Testimony." He argues that, if there is something wrong with moral testimony, it must be that acquiring a belief exclusively via moral testimony grants no moral understanding. When you settle the matter on testimonial grounds, writes Hopkins, "you have reasons for your belief but not moral reasons for it."[13] Hopkins suggests the following requirement on moral beliefs:

> The Requirement: having the right to a moral belief requires one to grasp the moral grounds for it.[14]

Hopkins's Requirement seems quite plausible. However, if it is indeed true, it blocks only the most outright form of moral deference, and does not block discorroboration. This is because, as I've argued above, *the information use of moral disagreement is both nonmoral and intellectually engaged*. When I acquire self-doubt through the observation of disagreement with a vegetarian, I am not acquiring a moral belief through testimony. The content of the beliefs I acquire through the discorroboration process are strictly about myself and the reliability of my own faculties. Any new belief I have acquired are strictly epistemic beliefs and don't fall under the scope of Hopkins's Requirement.

More important, *I fully understand the reasons for reducing my degree of confidence*. I can articulate the epistemic principles behind discorroboration. I understand the relationship between disagreement, discorroboration, and self-doubt. I can even give the reasons why I think my interlocutor a peer. Discorroboration escapes Hopkins's understanding requirement in precisely the same way that it escapes Wolff's autonomy requirement. In a discorroboration case, I both grasp the grounds I had for independently believing that p, and grasp the grounds for reducing my confidence in my independent reasoning that p.

2.2 THE DUTIES OF MORAL SELF-PERFECTION

I think, though, that further consideration of our intuitions about the importance of moral understanding will yield an important insight. I will argue that, if there is an active duty to achieve moral understanding, there must also be an active duty to seek out corroboration and discorroboration. Both duties are motivated by the same aim of moral self-improvement.

For this discussion, I would like to shift the focus from Hopkins's straightforward version of the requirement to what I take to be a more moderate, and more plausible, version of the understanding requirement, as developed by Philip Nickel.

In Hopkins's version, moral belief requires moral understanding on a case-by-case basis. In Nickel's view, an agent has *in general* a duty to understand his moral beliefs, thought his duty admits of occasional exceptions. I take Nickel to be largely correct about the issue, and will base my argument on his analysis.

Nickel's view is that:

> Morality aims at guiding action rationally, i.e. *from* a recognition of the relevant moral requirements. A moral agent must be responsive to morality as such . . .[15]

Part of what it is to act morally is to act from an understanding that the act is moral, says Nickel. The use of moral testimony typically provides correct moral belief without understanding. In these cases, "although moral testimony may give rise to a correct moral belief in these cases and thus 'work out' in the crudest sense, it does not provide a basis for morally good action," says Nickel.[16]

Nickel asks us to consider the following cases. First, imagine that my friend is addicted to heroin, and he asks for money. I tell him that I will lend him the money as soon as he gets his life together, and my justification is that my mother told me to say this.[17] Nickel says surely the friend should be unsatisfied, even if I defend my claim with explanations of my mother's greater age and experience. There is something, says Nickel, lacking about my justification if I offer no independent support for my moral beliefs, but only defer to my mother's judgment.

Next, imagine that a child is trained to act in accordance with certain rules, not by explanations, but by behavioral conditioning and punishment. Once the child is grown, she is unmoved by moral concepts. But she continues to perform moral-seeming actions (like returning borrowed objects) because she feels the force of her early conditioning.[18] Here, Nickel says that we have the intuition that something is missing from this action, which is that the action fails to come on the basis of recognition. Thus, says Nickel:

> It must be the case that morality requires one to *act* from an understanding of moral claims, and therefore to *have* an understanding of moral claims that are relevant to action.[19]

If our intuitions about these cases are true, says Nickel, then this Recognition Requirement must hold.

Nickel's version of the duty of moral understanding is significantly weaker than Hopkins's. Hopkins's thesis is that moral understanding is a requirement for moral action on a case-by-case basis. Nickel makes room for some special cases, in which a moral agent—who in general seeks and possesses moral understanding—occasionally steps aside and defers to testimony. Nickel here takes on board Karen Jones's argument that moral testimony can sometimes be useful as a corrective for bias. A moral agent who generally has moral understanding may sometimes acknowledge a blind spot in his own sensitivities. For example, says Jones, a morally sensitive man might still be insufficiently sensitive to some of the difficulties of a woman in the workplace. In a case where such sensitivity was vital, a morally sen-

sitive man may choose to defer to the judgment of somebody more sensitive to that particular matter.[20] Somebody who uses testimony occasionally and judiciously, as a corrective to bias, is still acting "with a conception of the relevant reasons in mind," says Nickel, even if they're unable to apply the relevant reasons accurately in each case. "It is only when one has a more global inability to grasp relevant reasons, or when one refuses to do so, that one cannot act morally well."[21] I take Nickel and Jones to be entirely correct in moderating the requirement, and this moderation will be crucial for my own argument.

Nickel claims that the Recognition Requirement is not only intuitive, but rational and reasonable. Understanding a claim, says Nickel, is vital to being able to apply the claim generally. When I depend on reliable moral testimony in one case, though I may get the action right in that single instance, I won't be able to reliably get similar cases right.[22] Second, says Nickel, if I don't understand the claims underneath my belief, my actions will be more error-prone. When I defer to moral testimony, doing the right thing requires that the situation remain static. Typically, I have to communicate the situation to my advisor, receive their advice, return to the situation, and then deploy that advice. Even if the advisor's testimony is correct, my acting correctly depends on the situation's remaining unchanged in its relevant details between receiving the testimony and acting. And, of course, lacking understanding, I don't know which parts of the situation are relevant. On the other hand, says Nickel, moral understanding allows me to respond flexibly to changing situations. If I know what's important, I'll be able to tell which changes to the situation are unimportant and which changes to the situations are highly relevant. An agent "must try to ensure that she is in a proper position to act from a recognition of what morality requires, and she is culpable for failing to do so." This shows that there is a "substantial epistemological duty to be able to tell what counts as a justification for a variety of moral claims," says Nickel.[23] If an agent doesn't strive for understanding, she will fail to be flexible and fail to retain her moral correctness in a fluid, ever-changing world. For Nickel, then, the connection between the duty for moral understanding and moral action is not constitutive, but pragmatic. Moral understanding is required for me to get it right more often.[24]

Nickel's view is, thus, that the duty of moral understanding is ultimately derived from a duty to perform the morally correct action.[25] Nickel's view fits very well with Jones's bias cases. The cases where Nickel allows a deviation from the duty for moral understanding are precisely those special cases where permitting an occasional deviation will lead to getting it right more often. An agent that possesses moral understanding in most cases can permit themselves a few deviations at the limits of their sensitivities, in those cases where doing so gets it right more often. But an agent that defers too often will fail to develop moral understanding, and so be unable to get it right in a very large number of cases.

Most important, says Nickel, those deviations are performed with a conception of the relevant reasons in mind. If I take myself to be less than perfectly sensitive to the subtle pressures that a woman faces in the workplace and defer my

judgment, I am still working with the right reasons in mind; I am still trying to act fairly and justly. I simply lack the ability to figure out the best way to achieve those goals in this particular situation, because I lack some relevant sensitivity.

For Nickel, the wrongs, demands, and duties associated with moral understanding gain their normative force as part of the means to getting it right. The duty here is *instrumental*—it derives its force from our duty to produce right action. This is why we are allowed to occasionally abandon the duty to right action, when it will help us act better. This is also why we cannot abandon that duty more than occasionally, for this would retard our own moral understanding, and so retard our ability to act well in an ever-changing world. Our moral duty to understand is not some stand-alone prerequisite for morality—if that were true, we would have no reason to occasionally circumvent the duty to understand. Our primary duty is to strive to get it right as much as we can, and this supports both the duty to understand in the majority of cases, and supports the occasional deviation from that duty in special cases, insofar as those deviations don't overly damage one's ability to act well. Furthermore, our duty to have moral understanding is a duty toward development—it is an active duty, that requires us to develop and perfect our abilities.

If the end of being morally correct is capable of generating one active epistemic duty, it should be capable of generating others. There is, after all, more than one factor that contributes to getting it right. Understanding certainly contributes to getting it right, but so does having properly functioning cognitive faculties. Forsaking moral understanding will certainly lead to frequent moral error. But having an error in one's cognitive process or a flaw in one's faculties of judgment will also lead to frequent error. Since moral correctness is a positive goal, it should motivate us to actively search out and eliminate any problems in our moral cognition. Discorroboration is one of the most significant forms of evidence for cognitive problems. Thus, the end of getting it right most often should give us an active duty to seek out other moral agents and use them to corroborate or discorroborate our moral cognition. Moral understanding and cognitive self-assessment are both substantial parts of the process of moral self-perfection.

My claims and Nickel's view fit so well because they are sensitive to the same fault-lines. When Nickel attends to Jones-style bias cases and allows for occasional deviations from moral understanding, he's taking into account the fact that we are fallible agents. Nickel and Jones are worried precisely about what I'm worried about: the need for a responsible moral agent to seek both understanding and engagement, and yet for a responsible moral agent to admit his own potential fallibilities, and move to make up for them with all tools available, including the moral judgments of others.

Understanding is a cognitive process. It is by its very nature aimed at getting things right. The cognitive values of reliability and truth call for many methods. If there is a positive duty for moral self-perfection, it ought to inspire us to develop our understanding. It ought also inspire us to use all means to check the reliability of our abilities and methods. Ignoring contrary testimony is surely a failure of my

moral epistemic duties as failing to achieve moral understanding. In fact, if the duties of moral self-improvement are as pervasive and substantial as Nickel suggests, then it seems like we may even have an active duty to regularly check our abilities, and so a duty to *seek out* contrary testimony and disagreement.

III. ALIENATION AND OBJECTIVITY

It may seem as if I'm asking agents to forsake their identity, to alienate themselves from their core values and commitments. But I've shown that the reasons and rules an agent uses in coming to self-doubt are reasons and rules that any rational agent would be committed to—part of the commitment to accuracy that is partially constitutive of rationality. The charge of alienation depends on the false presumption that our moral commitments are an essential part of our identity, but that our epistemic commitments are somehow alien. This cannot be the case for rational moral agents. Insofar as we take up rationality and its concomitant duties to accuracy, reliability, and truth, our epistemic commitments are also a crucial part of our identity.

Both the worries from autonomy and the worries from understanding are aimed at a particular sort of figure: an unthinking, obedient, passive agent who steps away from moral life, steps away from the responsibilities of being a rational, truth-seeking agent. But attending to discorroboration is the very opposite of being an unthinking, obedient agent. In using moral disagreement as a method for self-assessment, we are acting from reasons as phenomenally personal, endorsed, comprehensible, and directly felt as any moral reasons: we are acting from epistemic reasons and standards that apply to any cognitive project. We are acting from our commitment to truth, accuracy, and objectivity. The fact that these reasons point us toward evidence in the external world does not make them alien; any process aimed at getting right some objective, mind-independent properties ought to tune us toward the external world.

There is a difficulty to moral disagreement that is not there with empirical disagreement, I admit. This is because our moral beliefs are crucially *bivalent*; they are both personal and cognitive. They are important as both expressions of ourselves and as trackers of truth. This is why moral beliefs are the flash-point for this conflict. Moral beliefs, unlike everyday empirical beliefs, are simultaneously subject to very distinct rule-sets. As parts of our personality and identity, they are subject to considerations of self-preservation and self-expression. At the same time, our moral beliefs are important to us truth-trackers, and so they make themselves subject to the very impersonal rules and procedures of epistemic life. Any duties we have to be correct will call us not only to have an internal understanding of our moral beliefs, but to look outwards, to other people, to assess the reliability of our belief-formation process. But the impersonality of epistemic rules doesn't make them

alien to the agent. They are a crucial part of the agent—the part of the agent that is rational, that seeks to attune their beliefs to the objective, external world.

The worry that we are alienating ourselves through discorroboration comes from attending only to the personal valence of moral beliefs. But the pressure from discorroboration also comes from a place that is deeply embedded in any rational agent—the motive to get things right. I admit that the pervasive moral self-doubt that results attending to disagreement is difficult. It is painful. But it is not alienating.

NOTES

1. Hardwig (1991), Baier (1986), and Baier (1992).
2. There are also epistemic arguments for moral self-sufficiency, based on epistemic features common to all moral judgment. I discuss some such arguments in chapter 3 of Nguyen (2011).
3. The other significant argument excluding moral testimony on moral grounds is the argument from a special requirement for moral understanding, represented in the contemporary literature by Nickel (2001). That requirement is distinct from the autonomy requirement; I exclude discussion of it due to space constraints.
4. There is a slight quibble here: McGrath takes testimony and deference to be slightly different topics. For something to be testimony, it must be that another person actually speaks it. Her definition of deference does not require that another person actually speak his or her view; I can, for instance, come to think that my elder monk believes that one ought not kill insects on the basis of his constant painstaking actions to avoid killing insects. I take it that the difference is not important to my present line of inquiry, since my interest has always been in the epistemic weight of the contrary judgments of one's peers, with little attention to how those judgments are communicated.
5. McGrath (2011), 113.
6. Anscombe (1981), 44.
7. Wolff (1970), 13–14.
8. Wolff (1970), 13.
9. I am not claiming that there is a distinct line between bad obedience and the unproblematic cases. Rather, I am claiming that there is a continuum of cases, which can be best understood by analyzing the extremes.
10. And if we are not moral cognitivists, we shouldn't even be worried about moral disagreement in the first place.
11. The practical implications for moral humility depend on the existence of some principled relationship between the justifiability of action, and an agents' certainty in their reasons for undertaking the action. It seems to me very likely that there is such a relationship, but there is as yet no good account of the relationship. Developing such an account would be the next step in fully providing a theory of moral humility.
12. Either a Wolffian autonomy account or a moral understanding account is capable of independently explaining the wrong of moral deference. I suspect that the truth is that both accounts are correct, and that agents are subject both to requirements for autonomy and independent requirements to understand founded on epistemic norms.
13. Hopkins (2007), 19.
14. Hopkins (2007), 20.
15. Nickel (2001), 256.
16. Nickel (2001), 260
17. Nickel (2001), 256.

18. Nickel (2001), 257.
19. Nickel (2001), 257.
20. Jones (1999).
21. Nickel (2001), 264.
22. Nickel (2001), 261.
23. Nickel (2001), 261.
24. His account could even be used to show how acting independently would be better, even if in that particular case moral deference would lead to more morally correct action. This is because acting from one's own understanding is part of the process of developing one's understanding. Even if, in one's youth, one's parent is clearly more reliable than oneself, part of the process of becoming a morally able agent involves frequently striking out on one's own and working from one's own understanding. The errors one pays by avoiding deference in some cases will pay off in one's greater abilities down the line.

    In fact, it seems to me that Nickel has established something with these arguments slightly weaker than his initial claim. His language of his opening claim suggests that he will argue for a *constitutive* relationship between moral understanding and moral action, in general, where his arguments merely seem to establish that there is an *instrumental* relationship between moral understanding and moral action. My analysis of Nickel's position here relies on the content of his argument, rather than the language of those earlier passages.
25. This, presumably, is why this epistemological duty is not one that applies to all cognitive domains for all people, but does apply universally to the moral domain—only in the moral domain are we all subject to an active duty for correct performance.

REFERENCES

Ansombe, Elizabeth. 1981. *Ethics, Religion and Politics.* Minneapolis: University of Minnesota Press.

Baier, Annette. 1986. "Trust and Antitrust." *Ethics* 96 (2): 231–60.

Baier, Annette. 1992. "Trusting People." *Philosophical Perspectives* 6: 136–53.

Donagan, Alan. 1977. *The Theory of Morality.* Chicago: University of Chicago Press.

Foley, Richard. 2001. *Intellectual Trust in Oneself and Others.* Cambridge: Cambridge University Press.

Hopkins, Robert. 2007. "What Is Wrong with Moral Testimony?" *Philosophy and Phenomenological Research* 74 (3): 611–34.

Jones, Karen. 1999. "Second-hand Moral Knowledge." *Journal of Philosophy* 96 (2): 55–78.

McGrath, Sarah. 2011. "Skepticism about Moral Expertise as a Puzzle for Moral Realism." *Journal of Philosophy* 108 (3): 111–37.

Nguyen, C. Thi. 2011. "The Ethics of Uncertainty." PhD dissertation, University of California at Los Angeles.

Nickel, Philip. 2001. "Moral Testimony and Its Authority." *Ethical Theory and Moral Practice* 4 (3): 253–66.

Wolff, Robert Paul. 1970. *In Defense of Anarchism*. New York: Harper & Row.