Moral Understanding Between You and Me

Samuel Dishaw
UNC-Chapel Hill
sdishaw@unc.edu

[This is a penultimate draft, as of March 7, 2024, of a paper forthcoming in *Philosophy and Public Affairs.* Please cite the published version.]

Abstract: Much attention has been paid to moral understanding as an individual achievement, when a single agent gains insight into distinctly moral matters. Crucially overlooked, I argue, is the phenomenon of shared moral understanding, when you and I understand moral matters together, in a way that can't be reduced to each of us having moral understanding on our own. My argument pays close attention to two central moral practices: justifying our actions to others, and apologizing for wrongdoing. I argue that, whenever I owe you a justification or an apology, I thereby owe it to you to aim at our coming to a shared moral understanding. My argument has two upshots. The first is a novel explanation of the importance of moral understanding in our lives, one that emphasizes the importance of understanding moral reasons together. The second is a better understanding of the very obligations involved in two of our most central interpersonal moral practices.

Morality is, in large part, about how we should relate to other people. In the moral domain, it is possible to share an understanding of moral matters with the very person that understanding concerns. When you understand the reasons why it would be wrong for you to lie to your friend, you grasp something which your friend, too, could understand. This opens up a possibility: you might *share* a particular understanding of moral matters with the very person that understanding concerns. This possibility, in turn, raises a question: given that we can share moral understanding with others, ought we to? In what follows, I defend a positive answer to this question. In particular, I argue that two central moral phenomena—justifying our actions to others and apologizing for wrongdoing—always require shared moral understanding in order to be fully successful. Thus,

whenever I owe you a justification or an apology, I thereby owe it to you to try to bring it about that we share moral understanding.

One upshot of my argument is a novel explanation of the importance of moral understanding in our lives. Much attention has been paid to moral understanding as an individual achievement,[1] but the phenomenon of shared moral understanding has been widely overlooked. That omission is significant. If I am right, we can't fully explain why moral understanding matters by focusing exclusively on the moral understanding of individuals. What matters is also that we understand moral reasons *together*.

My argument also sheds light on the nature of two of our most central interpersonal moral practices. Justifying yourself to another person, I argue, constitutively involves aiming to share moral understanding with them about the reasons that justified your action. In a similar vein, I argue, an apology can be defective when it fails to reflect a shared moral understanding of the wrong done to its recipient. These claims, if correct, not only explain why we ought to aim at sharing moral understanding with others, they also illuminate what it takes to justify ourselves to others and to apologize for wrongdoing.

Since shared moral understanding has been widely overlooked, my first order of business in what follows is to give an account of what it takes for you and me to share moral understanding. I take up this task in sections 1 and 2. In section 3, I argue that shared moral understanding is the constitutive aim of interpersonal justification. In section 4, I argue that an apology should aim to reflect a shared moral understanding of the wrong it seeks to address. Section 5 concludes.

[1] Hopkins 2007; Hills 2009, 2015; Sliwa 2017; Callahan 2018; Howard 2018.

1. Individual Moral Understanding

My aim in this section and the next is to provide a working characterization of what it takes for moral agents to share moral understanding. On the account I provide, shared moral understanding is more than just the conjunction of two individuals having moral understanding. Still, it's a necessary condition on you and I sharing moral understanding that we each have the relevant moral understanding. So before I can say what it takes to share moral understanding, I need to say something about what it takes to have moral understanding in the first place.

Moral understanding, as its name would suggest, is understanding that is about the subject matter of morality. Although the exact boundaries of morality might not be perfectly clear, this won't be something we need to worry about. There are plenty of instances of understanding that are clearly about moral matters: understanding the wrongness of an action, what was disrespectful about a particular remark, why someone was justified in acting as they did, and so on and so forth.

The more important question for our purposes is what it takes to understand something, including moral matters. The first thing to point out is that understanding is a positive epistemic status. To understand why I ought to perform some action is for me be in a good epistemic standing vis-à-vis that particular subject matter. Importantly, understanding is a positive epistemic status that is graded along both objective and subjective dimensions. On the objective side of things, I understand something only if my conception of it is accurate. Thus, I understand why I ought to do A only if my beliefs about this subject matter are accurate: my belief that I ought to do A is true, and so is my belief about why I ought to do A—for example, that I ought to do A because I have promised you to do so. On the subjective side of things, I understand something only if I conceive of it in a way that makes sense to me. When I understand why I ought to do A, my belief

that I ought to do A is based directly in my own appreciation of the features that make it so that I ought to do A.[2] The structure of my thought mirrors the structure of the normative facts.

The subjective dimension of understanding has this important implication: you can know more than you understand. Not any collection of piecemeal moral knowledge amounts to genuine moral understanding. To illustrate, consider the following case:

> *Immoral Debt*
>
> Maya tells Jared that it was wrong for the IMF to make Madagascar pay its external debt towards developed countries because of the burdens it placed on the Malagasy population, and that the debt ought to have been abolished instead. This doesn't make any sense to Jared ("They borrowed the money," he thinks to himself. "Surely they *had* to pay it back?") But he trusts Maya's moral judgment and believes what she says.

By accepting Maya's testimony, Jared comes to know that Madagascar's external debt ought to have been abolished, as well as why it ought to have been abolished.[3] Yet Jared does not thereby come to understand why the debt ought to have been abolished. His belief that the debt ought to have been abolished is based not in his own appreciation of the first-order moral reasons at play (indeed, these reasons strike him as warranting the opposite conclusion), but rather in Maya's testimony. By contrast, Maya's belief that the debt ought to have been abolished is, we can imagine, based directly in the reasons that make that moral proposition true. Thus, whereas Maya

[2] Cf. Hopkins 2006; Hills 2009, 2016; Greco 2014; Grimm 2017. In principle, the subjective condition can be met independently of the accuracy condition: your belief that you ought to do A might be based directly in the features that you (mistakenly) take to explain why you ought to do A. Since I am interested in the mental state constituted by the unified conjunction of the two conditions, however, I will sometimes help myself to this factive gloss on the subjective condition.

[3] It is widely agreed that moral knowledge can be transmitted by testimony (see Jones 1999, Driver 2007, Hopkins 2007, Hills 2009, McGrath 2009, 2020, Sliwa 2017). I will not be calling that claim into question in this paper.

may understand why Madagascar's external debt ought to have been abolished, Jared does not. His belief that the debt ought to have been abolished is not based in his own appreciation of the reasons that make it so. He doesn't understand *himself* why that moral proposition is true.

To summarize: moral understanding is an epistemic achievement about moral matters. When you have moral understanding, you are getting things right: you represent moral reality in an accurate way. But in addition to this, you also represent that part of moral reality in a way that makes sense to you: your moral beliefs are themselves based directly in the reasons that make those beliefs true. Because moral understanding has this subjective dimension, you can come to know more, on the basis of moral testimony, than what you understand.[4]

2. Shared Moral Understanding

What has come before is largely familiar. We now turn to an overlooked but (I will argue) centrally important phenomenon: moral understanding that is shared between you and me.[5]

Shared moral understanding, as I will construe it, goes beyond merely having some bit of moral understanding *in common*. Having moral understanding in common is just a conjunctive fact: you and I have the same conception of why some action was wrong, and you and I both understand, on that basis, why that action was wrong. You have moral understanding, and so do I.

[4] A potential complication: won't this claim be rejected by those (e.g. Sliwa 2017) who think that understanding is ultimately reducible to knowledge? I don't think so. On Sliwa's view, an agent understands why p if and only if she has a sufficient amount of knowledge why p. This view leaves open the possibility that an agent might know enough about why p to know why p, without knowing enough to understand why p.

[5] As mentioned earlier, work on moral understanding has focused on the value and importance of *individual* moral understanding (see citations in fn. 1). One important exception, from the nearby literature on blame, is Fricker 2016.

When you and I *share* moral understanding, by contrast, not only do we have moral understanding in common, but the very fact of our having moral understanding in common is itself out in the open, in a shared space between you and me. We together acknowledge that we understand things the same way and in so doing we jointly commit to this way of understanding things.

To get a better handle on this notion of a shared space, consider the following example from Charles Taylor:

> Let's say that you and I are strangers travelling together through some southern country. It is terribly hot, the atmosphere is stifling. I turn to you and say: "Whew, it's hot." This doesn't tell you anything you didn't know; neither that it's hot, nor that I suffer from the heat. Both these facts were plain to you before […] What the expression has done here is to create a rapport between us […] Previously I knew that you were hot, and you know that I was hot, and I knew that you must know that I knew that … etc. … But now it's out there as a fact between *us* that it's stifling in here (1980, 294).

Taylor's example nicely illustrates a robust sense in which we can share a thought with another person, one that isn't reducible either to our having the same thought, or even to our knowing this about ourselves. As Taylor puts it, what is distinctive about the sort of public space that communication makes possible is that it creates a "common vantage point from which we survey the world together" (1980: 294).

When you and I share moral understanding, we survey moral considerations from such a common vantage point. In particular, when you and I share moral understanding, we jointly *commit*

to a way of understanding our moral reasons. This is because understanding, being belief-entailing, is itself a committal attitude. When you understand why an action is wrong, you hold a committal attitude: you regard that action as being wrong. When you and I share such an understanding, ours is a similarly committal attitude. Only now the content of that attitude is one to which we *jointly* commit, from a common vantage point, in understanding this part of moral reality together.[6]

Thus, the basic idea is that sharing moral understanding is structurally like sharing the experience of a hot day, with two small differences. One is that understanding is itself a committal attitude. The other is that what enters a shared space between us when we share moral understanding is usually more complex. When you and I share moral understanding about why some action was wrong, what figures in a shared space between you and me is a way of understanding why that action was wrong. We jointly commit not only to regarding this action as wrong, but also to a particular way of conceiving of what was wrong about it—we also commit, in other words, to treating certain considerations as the reasons why the action was wrong.

Such shared moral understanding, I will now argue, plays an essential role in our interpersonal moral practices. Within the central moral practices of interpersonal justification and apology, we owe it to others to aim at sharing moral understanding with them.

[6] This way of putting things departs from Margaret's Gilbert's treatment of Taylor's example, and from some of her views on joint commitment more generally (2013: 331 ff.). On Gilbert's view, our jointly committing to *x* is what *constitutes x*'s being part of a shared space between us. This view strikes me as too strong. You and I might be deliberating together about what to do tonight, and in the course of that joint deliberation considering whether to go out dancing tonight. We haven't decided to go out dancing, and yet it seems that the option of going out to dance is, in the relevant sense, in a shared space between you and me. So something can be in a shared space between you and me even if we haven't jointly committed to it yet. More generally, the notion of a shared space seems metaphysically prior to joint commitments: it's only *because* the fact that it's hot here is out in the open, in a shared space between you and me, that we *then* can jointly commit to recognizing that fact together. That said, the view on offer here—that shared moral understanding involves joint commitments—is of course compatible with either way of going on the issue of whether shared spaces are metaphysically prior to joint commitments, or vice-versa.

3. Interpersonal Justification

We should act towards others only in ways that we could justify to them. If I could not possibly justify breaking my promise to you, then it would be morally wrong for me to break that promise. Conversely, if it would be permissible for me to break my promise in this particular case, then it must be that I could, in principle, justify my doing so to you. We relate to others as we should when our actions have a certain property: the property of being justifiable to them.[7]

But that is not all. Relating to others as we should requires more than merely conducting ourselves in ways that we *could* justify to them. The act of justifying our actions to others is itself of moral significance. Beyond acting in ways that we could justify, we also, sometimes, owe it to others to in fact justify our actions to them. If I permissibly break my promise to you, it's not enough that I could, in principle, justify my doing so to you. Having broken my promise, I owe it to you to justify my doing so.

When do we owe it to others to justify our actions to them? I will not attempt to answer this question here. My own considered view is that there is a wide range of factors that can ground a duty to justify ourselves to others: having permissibly infringed someone's rights, being in a position of power, or simply being confronted by others about our actions when they have the standing to do so. I will not, however, assume any particular view about *when* we owe it to others to justify ourselves to them. The only assumption I make is that such a duty exists: there are instances in which we owe it to others to justify our actions to them.[8]

---

[7] This idea plays an especially central role within contractualism (Scanlon 1998), but it is also implicit in the Kantian thought that the law under which I act must be universally acceptable (Nagel 1995: 44ff.; Korsgaard 2018: 140).

[8] That such a duty exists is recognized, if only in passing, by a number of philosophers, including Montague 1988: 350; Elgin 2001: 105; and Sliwa 2019: 65, fn. 47. A lengthier discussion of interpersonal justification can be found in Frick 2016. Frick's primary focus in that essay, however, is interpersonal justification serving to persuade someone else that *they* ought to act in a certain way, which is not something we usually owe it to others to do. My concern in

This leaves us with a more fundamental question: *what* is it to justify your action to another person? In justifying ourselves to others, we are trying to achieve some end. So, one way to get at the nature of interpersonal justification is to ask ourselves what it aims at. That is, what is the constitutive aim of justifying your action to another person?

Let me try to make this question more precise. There are two complementary ways of thinking about constitutive aims. One is in terms of success conditions. If C is what it takes for some activity A to be successful, then C is the constitutive aim of A.[9] So, for example, if figuring out whether p is what it takes for inquiry into whether p to be successful, then figuring out whether p is the constitutive aim of inquiry into whether p. The second way of thinking about constitutive aims is in terms of what it takes for an agent to count as performing the activity at all.[10] Thus, if you're doing something that looks like inquiring into whether p, but you're not actually trying to figure out whether p, then you're not really inquiring into whether p. These two ways of thinking about constitutive aims are closely related. In order to count as performing some activity, you need to aim to do it successfully. So a constitutive aim is both the success condition of an activity and what it takes to count as doing that activity at all. There are thus two complementary ways to think about the constitutive aim of justifying your actions to others: in terms of its success conditions, and in terms of what you have to be trying to do in order to even count as justifying yourself to another person.

What, then, is the constitutive aim of justifying your action to another person? A very tempting answer is the following: to justify your action to someone is just to give them the reasons why your action is right. Call this the *Delivery Model* of interpersonal justification. The Delivery

---

this paper, by contrast, is the act of justifying one's *own* action to another person, which commonly enough is something that we owe to others.

[9] Velleman 2000: 246.

[10] Katsafanas 2018: 369-70.

Model provides a very neat account of interpersonal justification. It is also a very natural account: what *else* could it be to justify your action to another person, if not to give them the reasons why you were justified in acting as you did? [11]

And yet, we should reject the Delivery Model. Justifying yourself to another person is not just a matter of delivering some piece of information to them. As an account of interpersonal justification, the Delivery Model runs into two major problems.

First, the Delivery Model is unable to account for the fact that the duty to justify myself to you is specifically *mine* to carry out. When I owe you a justification, what I owe you is an action, not some piece of information: I owe it to you to justify myself as best I can. In other words, the content of the duty to provide a justification is best understood as making explicit reference to the agent who possesses that duty. This feature of the duty to justify oneself to another person is hard to make sense of on the Delivery Model. If the aim of interpersonal justification is just to deliver some information to you, it would seem to make no moral difference whether you receive that information from me rather than someone else. But this is mistaken. I can't delegate my duty to justify my actions to a moral expert in the way that I can delegate my duty to pay my taxes to a tax expert.

This point can be strengthened by considering the perspective of the person who is owed a justification. Suppose I pull out of a camping trip we were planning together. I owe it you to justify that decision. Suppose further that some well-informed third party takes it upon herself to give you the reasons why I was justified in pulling out of the camping trip, citing my many other commitments right now, my fear of bears, and so on. You might be glad to know about this, but it would nevertheless be reasonable for you to expect to hear these things from me. I am not entirely

[11] See Hills 2009: 106-108 and Wiland 2021: 65-67 for two discussions of justifying your actions to others (and of whether you need moral understanding to do so) that implicitly rely on the Delivery Model.

off the hook. If I never address the fact that I pulled out of the camping trip to you directly, knowing that our well-informed mutual acquaintance has provided you with adequate information already, I am making a moral mistake. You are entitled to more than just some piece of information: you are entitled to hearing that justification from me.

This is in contrast with other cases of directed moral duties that it seems perfectly fine for third parties to discharge on my behalf. Suppose I have borrowed a book from you that I promise to return to you by Sunday. On Friday, a mutual friend of ours, knowing that I can be forgetful and that I already finished reading the book, takes it upon himself to return the book to you. Although I did not fulfill my duty myself, our mutual friend has discharged my duty on my behalf: you now have the book in your possession, and there is nothing left for me to do. There is an important moral asymmetry, in other words, between the duty to provide a justification and the duty to return material objects. Because the Delivery Model reifies the aim of interpersonal justification, conceiving of it as a piece of information to be passed on to another person, it is unable to account for the personal nature of the duty to justify one's own actions to others, that is, the fact that the person who owes a justification and the person doing the justifying should be the same person.[12]

The second problem for the Delivery Model is independent of the first. Whereas the first problem for the Delivery Model concerns the role of the justifying agent, the second problem concerns the role of the addressee. Recall what the Delivery Model says about the aim of interpersonal justification: I successfully justify my action to you just in case I give you the reasons why my action was right. Note that, on this picture, it does not matter how you respond—whether you accept or reject the reasons I have provided, whether you call into question my way of describing the situation, whether you draw my attention to considerations you think I have

[12] This is not to say that it is never permissible to justify my action to you via an intermediary. In such cases, however, it matters that the intermediary is just that: a means of communication by why which I am addressing you.

overlooked, and so on. If the reasons I initially gave you were indeed the reasons why what I did was right, there is nothing further that I owe it to you to do. What you do with these reasons is your own business. My duty was just to give you a correct justification for my action, and I have done that already.

This, I take it, is an implausibly impoverished picture of what we are doing when we are justifying ourselves to others. In particular, what the Delivery Model fails to account for here is that the practice of interpersonal justification aims at a kind of uptake on the part of the addressee. What I am trying to do, in putting forward the reasons why I take myself to have been justified in acting as I did, is to justify my action *to you*. How you respond is the point of what I am doing. Whether I am successful in my attempt to justify myself to you depends, by my own lights, on whether I succeed in making the justification for my action apparent to you.

Generalizing further, it also seems that the Delivery Model is unable to make sense of why we engage in the practice of justifying ourselves to others. If I am completely indifferent as to whether or not you accept my justification for pulling out of the camping trip, why bother trying to justify myself to you? We engage in the practice of justifying our actions to others because we take their objections to our behavior seriously. If I take your objections seriously, I should also take your rejection of my justification seriously. Conversely, if I don't especially care whether or not you accept the justification I am putting forward for my action, it's unclear why I would be concerned with justifying my action to you in the first place.

Let us take stock. The Delivery Model has problems on both ends of the interpersonal relation. It loses sight of me, the justifying agent, by making it irrelevant who delivers the justification. It loses sight of you, the addressee of the justification, by making it irrelevant how you respond once the justification has been delivered. What we need, then, is an account that

addresses both of these shortcomings. Our account of interpersonal justification, that is, should explain the personal character of the duty to justify ourselves to others, as well as the importance of the response of the agent to whom such a justification is addressed.

With these desiderata on the table, I turn to a positive proposal. The constitutive aim of interpersonal justification, I will now argue, is shared moral understanding. When I am justifying myself to you, what I am doing is trying to bring it about that you understand my reasons for acting as I did, and thereby why I was justified in acting as I did. On this picture, interpersonal justification overlaps with the process of making oneself understood. I put my reasons out in the open, the reasons I take to justify my action, in the hope that you will understand me as having acted reasonably, or on good grounds. In so doing, I aim for us to understand the reasons for which I acted in the same light, as reasons justifying what I did. I aim, in other words, at shared moral understanding. Call this the *Shared Understanding Model* of interpersonal justification.

This statement of the view is somewhat abstract, so it will help to have an illustration of the basic idea that interpersonal justification aims at shared moral understanding. Consider the following case, which comes from Octavia Butler's dystopian novel, *The Parable of the Sower*. In it, the novel's fifteen-year-old protagonist, Lauren, lives within a tight-knit gated community precariously isolated from the lawless chaos on the outside. Lauren gets into trouble for trying to get the somewhat oblivious members of her community to realize that their gated paradise won't last much longer, and that they had better learn to survive outside the walls if they are to survive at all. Having confronted Lauren about this, her father tries to exact a promise from her: "No more scare talk" is the request. That's a no-go for Lauren: "I looked back at him, not wanting to give a promise that would be a lie." As far as Lauren is concerned, she's done nothing wrong; if anything, she takes herself to be morally required to alert her fellow community members about the dangers

they face. Since she can't accede to her father's request for a promise, she tries, instead, to justify herself to him. Her aim in doing so essentially involves uptake on his part: "I wanted him to know, to understand what I believed."[13] What Lauren believes, and what she thinks her father doesn't understand, is that their gated world is coming to an end, and that the people they both care about are completely unprepared for life outside the walls. And these facts, she also believes, amply justify telling her fellow community members how bad things are, even if it scares them stiff.

In justifying herself to her father, Lauren aims at shared moral understanding. She has a certain conception of the reasons why she is justified in telling people about the dangers they face, and the point of putting forward those reasons to her father is to bring it about that he share this understanding with her. The process of aiming at this shared moral understanding overlaps with the process of being understood. In justifying herself, Lauren aims for her father to understand *her*. But, crucially, not just *any* way of being understood would amount to success by Lauren's lights. Imagine that, in response to the reasons she provides, her father were to say to her something along the lines of: "Ah, I understand now why you thought all this scare talk was a good idea. But you're wrong about all this. There's no danger. So no more scare talk, okay?" This, rather obviously, is not the kind of uptake that Lauren is aiming at. Lauren doesn't merely want to be understood as acting for reasons that make sense from her perspective. She want to be understood as acting for *good* reasons, that is, for reasons that in fact justify her action ("It wasn't scare talk, we do need to learn what we can while there's still time").

In addition, it's plausible that Lauren aims at *sharing* moral understanding with her father, not just at having moral understanding in common. To see why, imagine that Lauren can tell, from the change in her father's tone of voice, or the deep furrow in his brow, that he has come to agree

[13] Butler 1993: 62.

that she did no wrong, and indeed that he understands the situation they are in and why she was justified in acting as she did in exactly the same way as Lauren does. But either out of stubbornness, fear or a misguided sense of parental authority, he refuses to recognize outwardly that Lauren acted rightly. Again it seems that something important is missing. It matters to Lauren that her father acknowledge that she acted for good reasons. It matters, in other words, that a certain way of understanding things is out in the open, in a shared space between them. This point is in fact closely related to the fact that Lauren aims to be understood as acting for good reasons. Indeed, to put things slightly differently, we might say that what Lauren aims at is to be understood as acting for reasons which, because they are good reasons, can be endorsed from a shared perspective that she and her father together occupy.[14]

My hope at this juncture is to have clarified what the Shared Understanding Model says, as well as made a case for thinking that it enjoys some initial plausibility. In addition, the Shared Understanding Model also provides a unified explanation of our two desiderata. First, it can explain the observation that, when I owe you a justification for my action, that duty is specifically mine to carry out. This feature of interpersonal justification can now be explained in terms of its aim. Other moral agents can give you a justification for my action, but only I can bring it about that you and I share moral understanding, by addressing you directly. Once we see that interpersonal justification aims at shared moral understanding, there's no more mystery as to why the person who owes a justification and the person doing the justifying should be the same person. Second, the view can also accommodate the observation that uptake matters to interpersonal justification—that whether I am successful in justifying myself to you depends on how you respond. This, in turn, enables the view to make sense of why we engage in the practice of

[14] See Korsgaard 1996 for the idea that normative reasons are necessarily reasons we can share with others.

justifying our actions to others in the first place. We engage in the practice of justifying our actions to others because, within our relationships with other moral agents, it matters to us whether we share moral understanding with them. Finally, the explanation of these two desiderata is *unified*. What explains the importance of *your* response to my reasons, as well the importance of *my* participation in providing these reasons, is simply that these are two aspects of the same goal: that of coming to a shared moral understanding.

I thus take it that the Shared Understanding Model has much going for it. Yet it is obviously not the only alternative to the Delivery Model. It will pay to consider two of these alternatives more closely. If they fail, their failure will provide, by way of elimination, some additional support for the Shared Understanding Model. Moreover, I will suggest that, in comparing the Shared Understanding Model to these two stronger alternatives, we can glean a further virtue of the account.

One such alternative view is that, when we justify ourselves to others, what we are doing is acknowledging another person's standing to issue a demand for justification. Following Darwall, we might conceive of these demands as being essentially second-personal. Your right to a justification is a "summons" addressed to me in particular.[15] The aim of interpersonal justification is to respond to this demand: to give you, personally, the justification to which you are entitled. Call this the *moral address* view. One virtue of this account is that, unlike the Delivery Model, it can easily explain why the duty to justify ourselves cannot be delegated to third-parties. Since the demand for justification is a second-personal demand made by you and addressed to me, no one can discharge this duty for me.

[15] Darwall 2006: 21 ff.

The moral address view is a clear improvement on the Delivery Model. However, it also inherits one of its flaws. This is because the moral address view is a version of the Delivery Model, one to which a personal touch has been added. The aim of interpersonal justification, on the moral address view, remains to deliver or provide a justification, with the caveat that this information must now be delivered in person, in response to a second-personal demand. Yet part of the problem with the Delivery Model was that, if I owe you a justification, I owe you more than the mere *delivery* of my reasons. I also owe it to you to reason with you, if, for instance, you disagree with my justification, or question my way of describing the events. The moral address view inherits this problem. The aim of interpersonal justification it provides is insufficiently sensitive to uptake from the person to whom such justification is owed. What should matter to me is not just whether I give you my reason *myself*, but also how *you* respond.

We can make this problem sharper by noting the following: in principle, I can respond to your second-personal demand for a justification, giving you in person the information to which you have a claim, while being completely *indifferent* to whether you think you have been treated rightly or not. I might say: 'Look, I'll tell you, for the record, why I was justified in doing what I did. But just so we're clear, I don't care one way or another what you think." A justification provided in this spirit of indifference towards you is defective. It is so not merely because I tell you your attitudes are of no concern to me. The indifference itself is what makes my justifying myself to you defective. I might also tell you my reason and walk away before you can respond. That way of delivering the justification would also express my indifference to your views about whether you have been treated rightly. It would still be a defective way of justifying myself to you. And yet both examples would count as responding to a second-personal demand made by you and addressed to me. In both cases, I have delivered my justification in person. Thus, the moral address

view is inadequate. The aim of interpersonal justification is not just to respond to a second-personal demand that I provide you with my reasons.

Where the moral address view goes wrong is in neglecting the uptake at which interpersonal justification aims. One might acknowledge this point and yet insist that the uptake at which interpersonal justification aims need not be anything as robust as shared moral understanding. For instance, one might think that the following is sufficient: I should aim for you to *know* that I was justified in acting as I did. This view avoids the problem of indifference outlined above. If I aim for you to know that I was justified in acting as I did, surely I am not indifferent to your actual views about whether you have been treated rightly. The knowledge view thus looks like a strong alternative to the Shared Understanding Model.

However, the knowledge view runs into a different problem. One way I can bring it about that you know some claim is by inviting you to take my word for it. But asking you to take my word for it that I was justified in acting as I did is something altogether very different from trying to justify my action to you. Suppose you entrusted some information to me in confidence, and you have now found out that I passed on this information to a mutual friend of ours. When you confront me about this, I respond by reassuring you that I did no wrong: you can take my word for it, I say, that in this particular case my breach of confidence was justified. In asking you to take my word for it that I did the right thing, I am aiming for you to know that I was justified in acting as I did. But I am not trying to justify myself to you. If anything, I am asking you not to demand a justification from me. So, knowledge is not the aim of interpersonal justification.

Against this line of reasoning, it may be urged that what is missing in the case above is that you don't know *why* I was justified in acting as I did. This suggests an improved version of the knowledge view: the aim of interpersonal justification is for you to know why I was justified in

acting as I did. This version of the view, however, remains equally vulnerable to the problem at hand. Suppose I tell you why I was justified in divulging the information to our mutual friend: they had a right to know about it. Suppose, further, that you are unconvinced: you disagree that our mutual friend had a right to know what you told me, or perhaps you deny that their having such a right licensed *me* in telling them about it. To overcome the impasse, I might ask you to take my word that the following is true: I was justified in divulging the information to our mutual friend, even though you had entrusted it to me in confidence, because they had a right to know what you told me. In asking you believe *that* claim on my authority, I might well be aiming for you to know why I was justified in acting as I did. But, again, I am clearly no longer trying to justify my action to you. If anything, I am asking you not to have to justify myself any further. So, the aim of interpersonal justification is not that you know why I was justified in acting as I did. More generally, it seems, the aim of interpersonal justification can't be knowledge, no matter its propositional *content*. Knowledge can always be acquired on the basis of someone else's authority, but trying to secure agreement by appeal to authority is antithetical to the practice of interpersonal justification. In justifying my action, I don't appeal to my authority. I appeal to your judgment.

Our examination of the moral address view and the knowledge view thus reveals, by way of contrast, a further virtue of the shared moral understanding model: on this model, the duty to justify ourselves requires relating to others as moral equals. One way we can fail to relate to others as moral equals, as in the moral address view, is by running roughshod over the moral attitudes of others—by being simply indifferent to what they make of our reasons. Another way we can fail to relate to others as moral equals, as in the knowledge view, is by caring about this uptake but trying to secure it by appealing the authority of our own moral judgment. By contrast, on the shared moral understanding model, I do not run roughshod over your own moral views. I do aim for you to

accept my justification. Yet I aim for you to accept my justification not on my authority, but rather on the basis of your own capacity to understand the reasons I take to have justified acting as I did. I relate to you as a moral equal.

This concludes my defense of the Shared Understanding Model. The upshot is the first argument for the main thesis of this paper. After all, we clearly sometimes owe it to others to justify our actions to them. But as I've just argued, shared moral understanding is the aim of interpersonal justification. Trying to bring about shared moral understanding is just part of what it is to try to justify ourselves to others. Thus, when I owe it to you to justify my action to you, what I owe you is in part the following: I owe it to you to aim at our coming to share moral understanding. This, in turn, is just an instance of the thesis I have set out to defend: that we have directed obligations to aim at shared moral understanding.

4. Apologizing

You can justify your action to someone else only if you acted rightly. But we do not always act in ways that are, even in principle, justifiable to others. Sometimes, we act in ways that we could not possibly justify to others. To treat someone in a way we could not possibly justify to them is to wrong them. When we've wronged someone, we shouldn't try to justify ourselves. We should apologize.[16] The duty to apologize is thus a derivative or secondary duty in the following sense: it comes into existence when a directed duty is violated, and the apology is owed to the very same person who was wronged when the original duty was flouted.

[16] For the sake of simplicity, I will be setting aside cases of justified wronging.

Apologies are part of the basic fabric of interpersonal morality. You are entitled to an apology if I flout a moral obligation I owe to you. A failure to apologize adds insult to injury. When no apology is forthcoming, you are warranted in resenting the absence of an apology itself, in addition to whatever behavior called for an apology in the first place.

What is perhaps less obvious is what the content of the duty to apologize amounts to: what do I owe to someone, when I owe them an apology? Equivalently: when I owe you an apology, what do you have a right to expect from me? Clearly what we owe by way of apology is more than just uttering the words "I'm sorry." You can say you are sorry without meaning it. Such an apology would be defective. What I owe you is a good apology. Our question in this section is what that amounts to.

In answering this question, we should keep in mind what apologies are for. The point of an apology is to repair the relationship between victim and offender, or at least to go some way towards repairing it.[17] The duty to apologize thus falls under the broader duty to repair or redress a moral wrong. The content of the duty to apologize should be sensitive to that aim. A good apology, the kind that we owe to another other, should be one that does work towards such moral repair.

My aim in this section will be to argue that apologizing has an epistemic dimension: an apology should aim to reflect a shared moral understanding of the wrong done to its recipient. Before I present my own view, however, it will be useful to consider a more familiar and less demanding account of the duty to apologize, and to see where that view goes wrong.

[17] For compelling defenses of this view, see Urban-Walker 2006; Radzik 2009; Martin 2010. I should note, however, that the relationship-repair framework is not uncontroversial (see Helmreich 2015: 85; Gardner 2018: 91-98; Bennett 2022). In what follows, I will not attempt to defend the idea that apologies aim at relationship-repair directly. Rather, I will argue that the best articulation of this idea will give a central role to the notion of shared moral understanding.

Indeed, the example of an insincere apology suggests a natural fix: it's not enough just to say you're sorry, you have to feel sorry too. More carefully, we might say that a good apology is one that is both *sincere* and *heartfelt*. The sincerity condition requires that, when you apologize for what you did, you sincerely believe that what you did was wrong. The heartfelt condition requires that you feel bad about what you did. Call this the *Feeling Bad Model* of apology. It says that a good apology is one that communicates that I feel (sufficiently) bad about what I did, where the nature of the bad feeling—guilt, remorse, contriteness—is whatever feeling is either constituted by or appropriate in light of the judgment that I did something wrong.[18]

One initial worry with the Feeling Bad Model is that it makes it too easy to delegate the duty to apologize to third parties. If the function of an apology is to communicate some information about my mental states—that I believe I acted wrongly, and that I feel bad about what I did—there seems to be no obstacle to my delegating that duty to a third party. This seems misguided: when I owe you an apology, you have a right to receiving that apology from me, not just from a delegate acting on my behalf. Similarly, if a mutual friend of ours takes it upon herself to tell you that I feel terrible about breaking my recent promise to you, I still owe you an apology. I'm not off the hook for apologizing to you, just because you have been apprised of how bad I feel about what I did.[19] It's unclear how to make sense of this on the Feeling Bad Model. If the purpose of an apology is to communicate to you that I feel bad about what I did, it seems that we can easily delegate the duty of apologizing for our actions, and that third parties can relieve us from the duty to apologize by making such an apology redundant. These two implications of the Feeling Bad Model fly in the

[18] The *locus classicus* of this view is Searle 1979: 4; 15. See also Darwall (2017: 42), who describes apology as guilt's "natural expression." I discuss a Darwallian version of the view below.
[19] Cf. Helmreich (2015: 78).

face of a basic truth about apology: when you owe an apology, you have to apologize yourself. Other people cannot apologize for you.

Defenders of the Feeling Bad Model might reply that I have misconstrued the communicative function of an apology. The function of an apology is not merely to communicate information about my mental states. Rather its function is to *express* how bad I feel about what I did, which is something only I can do. Other people can tell you about how I feel, but only I can express how I feel. Thus, on this version of the view, a good apology is one that expresses that I feel (sufficiently) bad about what I did, where the relevant notion of expression is essentially first-personal, in the sense that only my words and actions can be an expression of my feelings.

Even on this revised interpretation, however, the Feeling Bad Model still runs into a version of the redundancy problem. Just as a mutual friend of ours might, without my knowledge, tell you that I feel terrible about what I did, so too might *I*, unwittingly, express how bad I feel about what I did. My aggressive tone in response to your reproach, the way I grasp at excuses, or blame you in turn, may all very well be the expression of my feelings of *remorse*.[20] If so, then, in rejecting your reproach, and blaming you in turn, I will have succeeded, unwittingly, in expressing how bad I feel about what I did. But clearly I am not off the hook for apologizing to you. I still owe you an apology. So, there is more to apologizing than merely expressing a sufficient amount of bad feeling. Notice that this is just an extension of the basic point about duty delegation. Just as I am not off the hook when third parties communicate to you how bad I feel about what I did, so too am I not off the hook when I unwittingly express feelings of remorse. Just as other people cannot

[20] See Sliwa (forthcoming) for a discussion of an example along these lines.

discharge my duty to apologize on my behalf, so too I cannot discharge my duty to apologize on my own behalf through what I unwittingly do.[21]

A more promising way of strengthening the Feeling Bad Model is to turn to an analogue of a view discussed in the previous section: the moral address view. That is, one might suggest, that the duty to apologize, like the duty to justify ourselves to others, is a duty to respond to a second-personal demand. Your right to a sincere and heartfelt apology is a summons addressed to *me* in particular.[22] This view thus explains why others cannot release me from my duty to apologize by telling you about how bad I feel about what I did. In addition, the moral address view plausibly explains why I cannot satisfy my duty to apologize by unintentionally expressing my remorse. In order for me to acknowledge your right to a sincere and heartfelt apology, my communication of my guilt or regret must itself be intentional. It must be *addressed* to you rather than unwittingly conveyed. The moral address view is thus a clear improvement on the Feeling Bad Model. It adequately explains the personal character of the duty to apologize. At the same time, it also preserves the central idea of the Feeling Bad Model. On the moral address view, a successful apology is one that, personally addressed, conveys the extent to which guilty about what I did.

Because the moral address view remains a version of the Feeling Bad Model, it inherits a different problem common to all forms of this view: it fails to account for the importance of the apologizee's perspective on the wrong at issue. As the recipient of my apology, you have a right to expect more from me than just *my* feeling terrible about what I did.

---

[21] One might object that, in the example above, I haven't actually *said* "I'm sorry," and that for this reason I still owe you an apology. But that detail is not essential to the case. I might be that I in fact *apologize* in a hostile tone, or defensively ("I'm sorry, OK, what *more* do you want from me?"), and that this utterance is an expression of my remorse. In this case we would want to say that I still owe you a (proper) apology, even though I gave you an apology that expressed my remorse.

[22] Indeed, on Darwall's view, guilt itself is second-personal already: in feeling guilty, we recognize another person's authority to make a demand on us (2006: 72).

Here is a real-life example of the problem I have in mind. In June 2020, Drew Brees, quarterback for the New Orleans Saints and one of the public faces of the National Football League, was asked during an interview how he felt about players and teams kneeling during the national anthem before the start of games, as part of ongoing protests against racial injustice and police brutality. Brees answered, "I will never agree with anybody disrespecting the flag of the United States of America or our country," and went on to explain his view by describing how, during the national anthem, he envisions his grand-fathers and the sacrifices they made fighting in World War II. Backlash against his comments was predictably harsh and swift—so much so that less than twenty-four hours later, Brees issued a public video apology in which he apologized for his comments and the hurt they caused. One sports commentator, Michael Wilbon, took issue with the apology, and shared his misgivings about it in his show:

> I believe the apology. I believe it was as sincere and as heartfelt as he could possibly be and reflects that. But I'm angry today, Tony. Because even Drew Brees in his apology, he doesn't address what it was that ticked off so many people, including me, which is essentially the questioning of the patriotism of anybody—and let's say the "anybody" in this case is black folks—who want to take a knee, and have a protest during the anthem. […] Most white Americans feel that kneeling to the flag is disrespectful of the flag and is somewhat at odds with what their forefathers did, as if *our* forefathers weren't there dying on the same battlefield, which angers black folks, which angers me. […] Drew Brees' grandaddy came back home, perhaps to a local parade. My uncles and dad came back home to "Get on the back of the bus. No, you can't live in this neighborhood. No, you can't try the clothes in this department store." And they're all veterans. They all sacrificed and

> served this nation. Drew Brees still doesn't get that. At least the statement doesn't reflect that to me.

Wilbon's complaint is that Brees still doesn't understand what was offensive about his comments. This complaint seems perfectly legitimate. If Wilbon is right, this is a counterexample to the Feeling Bad Model, including its second-personal variant. After all, by Wilbon's own admission, Brees' apology was both *sincere* and *heartfelt*. It was, moreover, *addressed* as such to those whom the comments had offended. Despite this, the apology was defective. It fell short of what the recipients of the apology had a right to expect. More generally, what this example illustrates is a structural lacuna in the Feeling Bad Model. On the Feeling Bad Model (including its second-personal variant), what fundamentally matters to an apology is whether *I* feel sufficiently bad about what I did. As we have just seen, this leaves out something crucial. An apology is a response to a wrong done to *you*. Your perspective should matter. In apologizing to you, I should at least try to understand your reasons for objecting to my conduct.

The upshot of our example is thus not entirely negative. As Wilbon's response helpfully brings out, Brees' failure of moral understanding—a failure to understand why some moral fact obtained—is at the very same time a failure of interpersonal understanding—a failure to understand the point of view of other people. Brees' failure to understand what was offensive about his comments is, by the same token, a failure to understand what the people his comments angered were angry about. This suggests a natural way of remedying the structural problem noted above. The natural amendment is this: an apology should aim to reflect a shared moral understanding of the wrong done to its recipient. Call this the *Shared Understanding Condition* on apology.

We can think of the Shared Understanding Condition as a complex condition built up from more elementary ones. First, in apologizing to you, I need to understand what I did wrong. Second, I need to understand what I did wrong in relation to you, from your perspective: I need to recognize the appropriateness of your reasons for resenting[23] or otherwise objecting to what I did. Third, that very recognition of your reasons needs to be addressed to you: I need to acknowledge, to you, that your reasons for objecting to what I did are my reasons too.

To illustrate, suppose I was to drop off your portfolio at an art school you were hoping to enroll in. I somehow manage to forget your portfolio on the subway and miss the application deadline. Knowing how much this school means to you, and how upset you will be that I missed the deadline, I put off breaking the bad news to you for days. When I finally muster the courage to apologize to you, emphasizing how bad I feel about messing up the delivery, I am met with a surprise: although you are upset that your portfolio didn't make it to the admissions office in time, what upsets you the most is that I didn't trust you enough to tell you about it immediately. My apology misfires. It's not just that my apology reflects a failure to understand *some* respect in which my behavior was objectionable (though of course this is true). More significantly, what matters is that I failed to understand *your* legitimate reasons for objecting to what I did. I failed to recognize normative reasons that are yours.

The Shared Understanding Condition vindicates the idea that, when I owe you an apology, your thoughts and attitudes about the wrong done to you should matter to me. Part of my duty to apologize to you, on this view, is to try to understand your reasons and make them my own: to try

[23] In this section, I am using "resentment" in a somewhat loose sense that is meant to include other negative attitudes such as anger, disappointment, hurt feelings, and so on.

to share moral understanding with you. If I don't even try to understand your reasons for resenting what I did, this gives you a distinctive kind of warrant for rejecting my apology. [24]

The Shared Understanding Condition also explains why the duty to apologize cannot be easily delegated or unintentionally discharged. Other people can tell you that I feel bad about what I did, or even that I understand why what I did was wrong. But only I can make that moral understanding shared between us, by acknowledging your reasons for objecting to what I did. Moreover, since sharing moral understanding requires a joint commitment to the content of that understanding, sharing moral understanding is necessarily intentional. I cannot unintentionally commit to making your reasons my own.[25]

Finally, the Shared Understanding Condition provides a unified explanation of these two central features of apology. What explains why I must apologize *myself*, as well as why *your* perspective on the wrong matters, is one and the same fact: that an apology should try to bring about moral understanding that is shared between *us*.

I thus take it that the Shared Understanding Condition has a lot going in its favor. But we might still wonder why that norm is correct. That is, we might wonder *why*, in apologizing to you, I owe it to you to understand your reasons for resenting what I did.

To answer this question, it will help to have in view what apologies are for. The point of my apologizing to you, recall, is to repair the relationship between you and me. That relationship needs repairing because I wronged you: I acted in a way that showed insufficient regard for your

---

[24] This raises a question: what if you resent me for the wrong reasons, or blame me to an unreasonable extent? I consider this worry in the next section.

[25] An important corollary: the Shared Understanding Condition gives a straightforward explanation for the fact that apologizing on the basis of moral testimony is defective. I clearly cannot apologize to you simply because someone whose moral testimony I trust told me that I wronged you. If you realized that I believe I acted wrongly merely on someone else's say-so, that would, again, give you a special kind of warrant for rejecting my apology. I cannot simply act as a mouthpiece for someone else's moral understanding. My apology needs to reflect my own recognition of your reasons for resenting what I did.

reasons to object to my behaviour. In acting as I did, I failed to understand, or acknowledge, the force of your reasons to object to my action. But notice that your reasons for objecting to my action, ex ante, are the *very same reasons* that you have, ex post, for resenting what I have done. Just as you can object to my divulging a secret of yours on the grounds that I promised not to tell anyone, so too will you resent me for divulging your secret on those very same grounds. Thus, if I fail to understand your reasons for resenting what they did, I reiterate the very same moral mistake that is part of why the relationship between you and me is impaired in the first place. This, rather obviously, undermines the point my apology. It prevents it from serving its function as an instrumental of moral repair. To that extent, an apology that fails to reflect a shared moral understanding is defective. It falls short of what its recipient has a right to expect.

If this is correct, we have our second argument for the main thesis of this paper. We sometimes owe it to others to apologize to them. When we owe an apology, what we owe is an apology that reflects a shared moral understanding of the wrong done to its recipient. Thus, in the context of apologizing, we owe it to others share moral understanding with them, which is just another instance of the thesis I have set out to defend: that we have directed obligations to aim at shared moral understanding.

5. Reasoning with the Unreasonable

In closing, let me address what may be a lingering worry about the view I have defended so far. The claim that we have a duty to try to share moral understanding with others, a duty nested in the duty to justify ourselves and apologize, may in some cases seem to be too strong. What if my

interlocutor is unreasonable, or unwilling to reach a shared moral understanding with me? How could I be under a moral obligation even to *try* to bring it about that I share moral understanding with another person, when I know that my attempt to do so is all but certainly doomed to end in failure?

For instance, suppose I am, inexcusably, 5 minutes late to an appointment with you. I owe you an apology. You take my being late, however, as an egregious wrong, an affront to your honor, and you reject my apology on the grounds that it does not acknowledge this fact. In such a case, it seems, we would go too far in saying that I *owe* it to you to try to reach a shared moral understanding with you. A similar worry carries over to interpersonal justification. Suppose I am, this time, justifiably 5 minutes late, because I had to stop and help someone in a traffic accident. And yet suppose further that you reject even *that* reason as justifying my being 5 minutes late. Again it seems implausible to claim that I am under a moral obligation to try to reach a shared moral understanding with you.[26]

Note that the worry cannot be avoided simply by denying that we have a duty to apologize or justify ourselves in such cases. An apology and a justification *are* owed in the two examples above. And yet, it seems, there is no duty to try to reach a shared moral understanding. What these cases seem to illustrate, then, is that we can owe a justification or apology to others *without* being under an obligation even to try to reach a shared moral understanding with them. These cases thus pose a serious challenge to the view I have defended so far, which says that the duty to justify ourselves and to apologize includes within it, and so entails, a duty to try to share moral understanding with others.[27]

---

[26] I am indebted to an associate editor of this journal for suggesting this particular way of pressing the worry, and for these two examples that vividly illustrate it.

[27] This worry is made even more pressing by the fact that the view's strongest rival, the moral address view, does not face it at all.

To meet this challenge, we need to make two general observations about the view I have put forward. The first observation is that the duty to justify ourselves to others, or to apologize, is a *directed* duty: it is the logical flipside of another person's claim or right against me that I act in a certain way.[28] Such claims have an interesting feature: a claim-holder can *forfeit* their claims through their own actions. For instance, suppose I promise to play tennis with you over the weekend. It's now Thursday and I reach out to you to set a time for us to play. You don't respond. Friday comes and goes, and so does Saturday. Come Sunday, I *could* ring your doorbell, track you down, to make sure I keep my promise. But I am not under a duty to do so. Though keeping my promise is something I could do, you have, through your own uncooperativeness forfeited your *claim* against me that we play tennis together this weekend.

The second observation concerns the structure of the right to a justification or to an apology. The view I have defended so far pairs naturally with the idea that these rights are best construed as a *cluster* of rights.[29] To say in this context that the right to a justification or an apology is really a cluster of rights is to say that it includes rights to the performance of separate actions. For instance, your right that I justify myself to you includes, minimally, a claim that I give you the justification for my action, but it also includes, or so I have argued, a right against me that I reason with you with the aim of our coming to a shared moral understanding. Similarly, your right to an apology includes, minimally, a right that that I tell you how sorry I am about what I did, but it also includes, or so I have argued, a right against me that I try to understand your reasons for resenting what I did.

Together, these two observations provide an answer to the objection raised above. We can now say that those who are unreasonable towards us forfeit some, but not all, of the rights included

[28] Thomson 1990; Wallace 2020.
[29] See Thomson 1990: 67.

in their right to a justification or an apology. If you take my being 5 minutes late as an egregious affront to your honor and reject my apology on those grounds, you forfeit your right against me that I try to reach a shared moral understanding with you. Similarly, if you reject my justification that I was 5 minutes late because I helped someone in a traffic accident, you plausibly forfeit your claim that I reason with you about whether that consideration really does justify my conduct. What explains why an unreasonable interlocutor forfeits these rights in such cases is that they are uncooperative. Through their uncooperativeness, they forfeit their claim against our engaging in a cooperative activity with them: that of reasoning with them, with the aim of reaching a shared moral understanding. This conclusion is thus consistent with thinking that the unreasonable retain some rights against us. After all, giving another person your justification, or simply saying 'I'm sorry,' does not require any cooperation from them beyond their willingness to receive that information. If a person is cooperative to this minimal extent, they retain their claim to some of what their right to a justification or apology entitles them, even if they may have forfeited other rights within that cluster.

This general picture is plausible beyond its ability to respond to the challenge of unreasonable interlocutors. One virtue of this picture is that it can explain the discretionary leeway of those who are owed a justification or an apology. Perhaps I have slighted you in a way I don't quite understand. It would be good if I understood what, exactly, offended you, but you would prefer we move on, so you say: "I don't need you to understand. Just don't do it again." Here, you *waive* your claim against me that I try to understand your reasons for objecting to my conduct. I cease to have a duty to try to reach a shared moral understanding with you, not because you are unreasonable, but simply because you have released me from that duty. For this idea to make sense, however, it must have been that I was under such a duty to begin with—that you had such a right

to waive. A straightforward explanation of this discretionary power is provided by the logical flipside of the thesis defended in this paper: that the right to a justification or an apology includes, as one of its parts, a right that others try to reach a shared moral understanding with us.

In short, the duty to justify ourselves to others, or to apologize, is best construed as a cluster of duties, to which a cluster of claim-rights corresponds. Central among these rights is a right that those who owe us a justification or an apology try to reach a shared moral understanding with us. When a claim-holder is sufficiently unreasonable, however, they may thereby forfeit their claim that others try to reach a shared moral understanding with them, even if they retain the more minimal claim to be given a justification or offered an apology.

A further virtue of this general picture, as against an account such as the moral address view, is that it avoids providing an identical account of what we owe to the reasonable and the unreasonable. Morality, on the present account, requires something of us in our relations with the unreasonable, but not as much as it would were it not for their unreasonableness. This, too, seems correct. We owe something more, it seems, to those who cooperate with us in finding reasonable terms on which to live, than we do to those who refuse to do so. The idea that the unreasonable forfeit some but not all of their rights allows us to maintain this asymmetry.

I end with a cautionary note. The view that the unreasonable forfeit certain claims against us must be tempered by a careful restriction of the kinds of unreasonableness that result in such forfeiture, and by a corresponding reluctance to regard others as simply being unreasonable. For instance, a person's *past* unreasonable conduct does not, I believe, result in their lacking a claim that I try to reach a shared moral understanding with them on a present occasion. I am not off the hook to try to come to a shared moral understanding with you just because you have unreasonably rejected my justifications in the past. If I and others no longer saw any point in reasoning with you,

how could you *earn back* the right to be reasoned with? After all, it is through the mutual exchange of reasons that we reveal ourselves to be reasonable (or not). The idea that a person's track record might disqualify them from the full-blown version of our moral practices thus poses a significant obstacle to their rehabilitation within them.[30] Instead, it seems more plausible that a person maintains their right that we try to reach a shared moral understanding with them, on any given occasion, until they conduct themselves unreasonably on that occasion. Respecting others as moral equals requires viewing them as capable of responding to reasons, past failures notwithstanding.

For similar reasons, a person does not forfeit their claim to be reasoned with simply in virtue of the attitudes they *presently* hold, prior to one's offering a justification or apology. Perhaps I know, prior to apologizing to you, that you are furious and blame me out of proportion to what I did wrong. These attitudes alone do not let me off the hook. I do not get to hold your attitudes as a given and think to myself: 'coming to a shared moral understanding is off the table here, given how angry this person is. I will simply say I am sorry and move on.' Holding your current attitudes as a fixed fact about you fails to acknowledge your capacity to change your attitudes in response to reasons. You have not forfeited your claim that I try to come to a shared moral understanding with you simply because you currently blame me out of proportion to what I did wrong.

What these two observations suggest is that the kind of unreasonableness which results in a person forfeiting their claim to be reasoned with is not the unreasonableness of past conduct, nor the unreasonableness of presently held attitudes, but rather must be a form or intransigence or uncooperativeness with our actual attempt to reason with them. In our earlier example, for instance, what alters the nature of my duties to you is that you actually reject, unreasonably, my justification that I was late because I had to stop and help someone in a traffic accident. By unreasonably

[30] This point is indebted to an observation made by Shiffrin (2008: 489) in her discussion of promissory obligations.

rejecting that justification, you fail to do your part towards our coming to a shared moral understanding. Then—and only then—do I cease to be under a duty to do my mine: to further reason with you, with the aim of reaching such shared understanding.[31] Cases in which we owe a justification or an apology to unreasonable people will not thus, generally speaking, be cases in which a duty to try to reach a shared moral understanding is entirely absent. Rather, these will be cases in which we cease to be under a duty to try to reach a shared moral understanding with another person, having tried, initially, to do just that. Even in our relations with the unreasonable, then, the point of apologizing or justifying ourselves to others still remains to share moral understanding with them. The question, in such cases, is not so much whether we should try to come to a shared moral understanding, but rather how hard we must try.

6. Conclusion

In this paper, I argued for the importance of a kind of understanding about moral matters that essentially involves moral agents other than oneself: moral understanding that is shared with another person. In particular, I provided two arguments in support of the claim that we sometimes owe it to others to try to share moral understanding with them, the first based on the nature of interpersonal justification, the second on the norms of apology.

[31] Such a case thus differs from one in which you and I reasonably disagree about whether my conduct was justified. In cases of reasonable disagreement, we may both do our part towards reaching a shared moral understanding yet fail to find this common ground. The normative difference in such a case is that you do not forfeit any rights when you reject my justification. Assuming our disagreement is reasonable, it must be reasonable for you to initially reject my justification. And so I would, at that point, remain under a duty to do my part towards reaching a shared moral understanding with you, by further reasoning with you about whether that justification really was sufficient in the case at hand.

Although I presented these arguments separately, I also think that it is not a coincidence that they support the same conclusion. Justifying and apologizing are, at bottom, two sides of the same coin, each practice embodying the basic idea that moral reasons are essentially shareable. If I had sufficient moral reasons for doing what I did, it must be that you too could accept these reasons as justifying what I did. If you had decisive reasons to object to my conduct, it must be that I too could, and should, have taken those reasons as ruling out acting as I did. Moral reasons have a potential for uptake by others that is part of their very nature. Justifying and apologizing are ways that we actualize the potential of moral reasons to be shared between us.

According to a venerable tradition in moral philosophy, what it is to relate to others in the right way is to relate to them in ways that they could rationally accept (or not reasonably reject). That is, we relate to others in the right way when we act for reasons that have a distinctive *potential*: the potential to be the object of rational agreement between ourselves and others. For me to relate to you in the right way is for me to act on reasons that you and I could share.[32] But as we seen, relating to others as we should occasionally involves more than this. When we owe a justification or an apology, relating well to others also requires trying to *actualize* the potential for rational agreement inherent in moral reasons. Within these moral practices, we owe it to one another not only to act on reasons which we could in principle share, but also to try to bring it about that we in fact share these reasons, by understanding them together.

[32] For Kantians, being motivated to relate to others in the right way might involve thinking about what laws one would adopt if one were legislating in the kingdom of ends (Hill 1989), or what principles what be acceptable from anyone's point of view (Nagel 1995; Korsgaard 2018). For contractualists, this might involve thinking about what principles others could not reasonably reject (Scanlon 1998).

References

Bennett, C. (2022). What Goes On When We Apologize? *Journal of Ethics and Social Philosophy* 23 (1): 115-135.

Bengson, J. (2015). A Noetic Theory of Understanding and Intuition as Sense-Maker. *Inquiry* 58 (7-8): 495-532.

Butler, O. (1993). *The Parable of the Sower*. Grand Central Publishing.

Callahan, L.F. (2018). Moral Testimony: A Re-Conceived Understanding Explanation. *The Philosophical Quarterly* 68: 437-459.

Cohen, J. (1989). Belief and Acceptance. *Mind* 98: 367-389.

Darwall, S. (2006). *The Second-Person Standpoint*. Harvard University Press.

------ (2017). Trust as a Second-Personal Attitude (of the Heart), in Faulkner, P. (ed.) *The Philosophy of Trust*: 35-50. Oxford University Press.

Driver, J. (2007). Autonomy and the Asymmetry Problem for Moral Expertise. *Philosophical Studies* 128 (3): 619-644.

Elgin, C. (2001). Word Giving, Word Taking, in Byrne, A., Stalnaker, R. and Wedgwood, R. (eds.) *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*. MIT Press.

Howard, N. (2018). Sentimentalism about Moral Understanding. *Ethical Theory and Moral Practice* 21 (5): 1065-1078.

Frick, J. (2016). What We Owe to Hypocrites: Contractualism and the Speaker-Relativity of Justification. *Philosophy and Public Affairs* 44 (4): 223-265.

Fricker, M. (2016). What's the Point of Blame? A Paradigm Based Explanation. *Noûs* 50 (1): 165-183.

Gardner, J. (2018). From Personal Life to Private Law. Oxford University Press.

Gilbert, M. (2013). *Joint Commitment: How We Make the Social World*. Oxford University Press.

Greco, J. (2014). Episteme: Knowledge and Understanding, in Timpe, K. and Boyd, C. (eds.) *Virtues and their Vices*: 284-301. Oxford University Press.

Grimm, S. (2017). Understanding and Transparency, in Grimm, S., Baumberger, C., and Ammon, A. (eds.) *Explaining Understanding: New Perspective from Epistemology and Philosophy of Science*: 212-229. Routledge.

Helmreich, J. (2015). The Apologetic Stance. *Philosophy and Public Affairs* 43 (2): 75-108.

Hill, T.E. (1989). Kantian Constructivism in Ethics. *Ethics* 99 (4): 752-770.

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

------ (2015). The Intellectuals and the Virtues. *Ethics* 126 (1): 7-36.

------ (2016). Understanding Why. *Noûs* 50 (4): 661-688.

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

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

Katsafanas, P. (2018). Constitutivism about Practical Reasons, in Star, D. (ed.) *The Oxford Handbook of Reasons and Normativity*: 367-394. Oxford University Press.

Korsgaard, C. (1996). The Reasons we can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values, in *Creating the Kingdom of Ends*. Cambridge University Press.

------ (2018). *Fellow Creatures: Our Obligations to the Other Animals.* Oxford University Press.

Martin, A. (2010). Owning Up and Lowering Down: The Power of Apology. *The Journal of Philosophy* 107 (10): 534-553.

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

------ (2020). *Moral Knowledge*. Oxford University Press.

Montague, P. (1988). When Rights are Permissibly Infringed.

Nagel, T. (1995). *Equality and Partiality*. Oxford University Press.

Radzik, L. (2009). *Making Amends: Atonement in Morality, Law and Politics*. Oxford University Press.

Scanlon, T.M. (1998). *What We Owe to Each Other*. Harvard University Press.

Searle, J. (1979). *Expression and Meaning*. Cambridge University Press.

Shiffrin, S. (2008). Promising, Intimate Relationships, and Conventionalism. *Philosophical Review* 117 (4): 481-524.

Sinclair, T. (MS). Claims as Inputs.

Sliwa, P. (2017). Moral Understanding as Knowing Right from Wrong. *Ethics* 127 (3): 521-552.

------ (2019). The Power of Excuses. *Philosophy and Public Affairs* 47 (1): 37-71.

------ (forthcoming). Taking Responsibility, in Chang, R. (ed.) *Normative Philosophy: Conversations in Moral, Legal, and Political Philosophy*. Oxford University Press.

Stalnaker, R. (1984). *Inquiry*. MIT Press.

Taylor, C. (1980). Theories of Meaning. *Man and World* 13 (3-4): 281-302.

Thomson, J. J. (1990). *The Realm of Rights*. Harvard University Press.

Velleman, J. D. (2000). *The Possibility of Practical Reason*. Oxford University Press.

Urban-Walker, M. (2006). *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. Cambridge University Press.

Wallace, R.J. (2020). *The Moral Nexus*. Princeton University Press.

Wiland, E. (2021). *Guided by Voices: Moral Testimony, Advice and the Forging of a "We."* Oxford University Press.