

---

# MODIFICATION OF THE REACTIVE ATTITUDES

BY

DAVID GOLDMAN

---

**Abstract:** In 'Freedom and Resentment' P. F. Strawson argues that reactive attitudes like resentment and indignation cannot be eliminated altogether, because doing so would involve exiting interpersonal relationships altogether. I describe an alternative to resentment: a form of moral sadness about wrongdoing that, I argue, preserves our participation in interpersonal relationships. Substituting this moral sadness for resentment and indignation would amount to a deep and far-reaching change in the way we relate to each other – while keeping in place the interpersonal relationships, which, Strawson rightfully believes, cannot be eliminated.

In the 20th century, Mahatma Gandhi and Dr. Martin Luther King Jr. rejected antagonistic response to wrongdoing categorically. 'If [peace] is to be achieved,' King said in his Nobel Peace Prize Speech in 1964, 'man must evolve for all human conflict a method which rejects revenge, aggression, and retaliation. The foundation of such a method is love.'<sup>1</sup> Call what is being suggested here the Gandhi-King ideal. On this ideal, it would be better to eliminate antagonistic reactions to wrongdoing if we could. Even if we cannot eliminate such reactions entirely, we should condemn and disavow them whenever they do appear.

In this article I consider P. F. Strawson's defense of our moral 'concepts and practices' in his immensely influential essay 'Freedom and Resentment.'<sup>2</sup> Although Strawson does not endorse *revenge* or *retaliation*, he does argue that broadly antagonistic responses to wrongdoing are an important and ineliminable part of our moral lives.

Strawson situates our moral concepts and practices within *engaged interpersonal relationships*. It is an essential part of being in such relationships, he thinks, that we make certain broadly moral *demands* on each other. It is, in turn, essential to making a demand that we be susceptible to certain *attitudinal responses* if the demand should be violated. If Strawson

is correct, these relationships, demands, and responses stand or fall together: eliminating our susceptibility to certain attitudes in response to wrongdoing amounts to dropping the moral demand, and thereby exiting an interpersonal relationship.<sup>3</sup>

'Exiting an interpersonal relationship' might call to mind severing one's contact with someone. Strawson has in mind a much deeper and more costly change to the way one relates to a person. Exiting an interpersonal relationship involves adopting what Strawson calls the 'objective attitude' towards the person: one comes to view her as an object to be 'managed or handled or cured or trained,'<sup>4</sup> and ceases to relate to her as a fellow human being.

Resentment is Strawson's central example of the attitudinal responses that are essential to participation in interpersonal relationships. Resentment, as Strawson understands it, is antagonistic in at least two ways. First, it is connected (though sometimes indirectly connected) to motivations to do harm: resentment entails, Strawson says, a 'modification . . . of the general demand that another should, if possible, be spared suffering,' which modification can, sometimes, give rise to a 'preparedness to acquiesce in the infliction of suffering'<sup>5</sup> on a wrongdoer.

In addition, resentment seems to be a form of anger; and anger involves an affect which is, itself, antagonistic. (To see this, compare the experience of being angry with someone to the experience of being happy with someone.) Strawson himself does not claim that resentment is a form of anger; but that does seem to be a plausible elaboration on his view. At the very least, resentment often does involve an affective element resembling the affect of anger.

I think the connections Strawson sees between relationships, demands, and responses are real and important. But I do not think that specifically antagonistic attitudes must be a part of this close-knit group. In this article I develop an account of the important role that antagonistic responses do in fact play in our engaged, demanding interpersonal relationships. I then identify a form of sadness that can play the very same role without being antagonistic in any of the ways just sketched. If my arguments are sound, I will have eliminated one important reason<sup>6</sup> for thinking that the Gandhi-King ideal of eliminating antagonistic response must come at the high price of ceasing to demand goodwill and exiting interpersonal relationships.

It is worth pausing to note the relation of my argument to one made by Derk Pereboom in *Living Without Free Will*. In that book Pereboom argues for hard incompatibilism, which is 'the view that there is no freedom of the sort required for moral responsibility' (p. xxiii). Because hard incompatibilism is true, Pereboom thinks, reactions to wrongdoing like blame and resentment are inappropriate. After arguing for hard incompatibilism, Pereboom considers Strawson's claims that our engaged

interpersonal relationships require reactive attitudes like blame and resentment. He rejects these claims, arguing that engaged relationships can survive even without resentment and suggesting that resentment could be replaced by a 'kind of moral sadness . . . that would not be undermined by a belief in determinism.' Pereboom says of this sadness, 'I suspect that it can play much of the role that resentment and indignation more typically have in human relationships.'<sup>7</sup>

My project in this article dovetails with Pereboom's in that I, too, aim to describe a form of moral sadness that can play the same role typically played by resentment and indignation. But I am motivated to identify an alternative to resentment by a very different set of considerations. Pereboom wants to identify an alternative to resentment that could serve (much of) resentment's function in a world in which nobody is morally responsible for anything they do. I do not agree with Pereboom that nobody is morally responsible (though this is not the place to respond to his arguments). I want to identify an alternative to resentment that would not be antagonistic and so would not run afoul of the deep moral skepticism about antagonism expressed by King and Gandhi.<sup>8</sup>

In arguing that sadness could play the same role as resentment and indignation in our interpersonal lives, Pereboom focuses on the important fact that such reactions communicate an emotional and deeply felt distress over wrongdoing.<sup>9</sup> He argues that moral sadness can play the same role. I agree that resentment and the like play this important communicative role; but I do not think that fully captures their significance in our interpersonal relationships. I provide a more detailed account of resentment's role in our interpersonal lives in Section 2 below. In Section 3, drawing on the notion of a 'sharpening' of a basic emotion recently developed by Justin D'Arms and Daniel Jacobson,<sup>10</sup> I provide a more detailed specification of the sadness that, I argue, could fill the very same role in our interpersonal lives.<sup>11</sup>

## 1.

I will begin by describing the broad outlines of Strawson's argument. Strawson re-poses questions about responsibility and blame in terms of certain attitudes we have in response to wrongdoing: personal, reactive attitudes, when we are wronged ourselves, and impersonal, vicarious analogues of the reactive attitudes, when we consider wrongs directed at others.<sup>12</sup> When I am wronged, I will naturally feel resentful; when I see someone else wronged, I will feel indignant.

Strawson introduces these attitudes by describing the broader context within which they arise. He begins by pointing to what he calls a 'common-place': 'the very great importance that we attach to the attitudes and

intentions towards us of other human beings, and the great extent to which our personal feelings and reactions depend upon, or involve, our beliefs about those attitudes and intentions' (p. 75). This concern about the attitudes of others arises, Strawson says, within a wide variety of interpersonal relationships:

We should think of the many different kinds of relationship which we can have with other people – as sharers of a common interest; as members of the same family; as colleagues; as friends; as lovers; as chance parties to an enormous range of transactions and encounters (p. 76).

Within these relationships, Strawson continues, we demand goodwill or interpersonal regard of each other:

Then we should think, in each of these connections in turn, and in others, of the kind of importance we attach to the attitudes and intentions towards us of those who stand in these relationships to us . . . In general, we demand some degree of goodwill or regard on the part of those who stand in these relationships to us, though the forms we require it to take vary widely in different connections (p. 76).

Within any human interpersonal relationship, Strawson thinks, we demand goodwill or interpersonal regard. Resentment arises in reaction to violations of that demand.

Strawson expresses hope that focus on resentment and these related phenomena will generate new insight into old problems about moral responsibility. In particular, questions about whether determinism would, or should, undermine blame or responsibility can be re-posed as questions about whether determinism would, or should, undermine the resentment that arises within this context, or the attitudes which are its vicarious analogues.

Strawson focuses on determinism's impact on these reactive attitudes; but, he says, he does not know what the thesis of determinism is. He only knows that if the thesis of determinism is true, it is universally true: all behavior is determined.<sup>13</sup> If determinism undermines resentment and its interpersonal analogues, then, it does so in *every* case – it does so universally. Strawson aims to show that resentment could not be undermined universally. I am interested in this strong claim about resentment's resilience, not the threat posed by determinism in particular; so in what follows I set determinism to one side.

To argue that resentment could not be undermined universally, Strawson first maps out the ways in which resentment and the like are ordinarily undermined. He asks 'what sorts of special considerations might be expected to modify or mollify this feeling [of resentment] or remove it altogether' (p. 77).

Strawson's answer to this question draws on his understanding of resentment as a reaction to a violation of the demand for goodwill. Two sorts of considerations will inhibit resentment, so understood.

One type of inhibiting consideration shows that, contrary to initial appearances, the demand for interpersonal regard has been met. In the face of the 'appearance of this demand's being flouted or unfulfilled,' considerations of this type 'show this appearance to be mere appearance, and hence inhibit resentment' (p. 85). This sort of consideration inhibits a reaction to the violation of a demand, by showing that the demand was not violated in the first place. Strawson lists considerations like 'He didn't mean to' and 'He couldn't help it' (p. 77) as examples of this type of consideration.

Another type of inhibiting consideration undermines the demand for interpersonal regard itself. Such a consideration 'tends to inhibit resentment because it tends to inhibit ordinary inter-personal attitudes in general, and the kind of demand and expectation [i.e., for inter-personal regard] which those attitudes involve' (p. 86). This sort of consideration inhibits a reaction to a violation of a demand by *undermining the demand*. These considerations often indicate that the purported agent was not in fact an agent with a will, capable of showing interpersonal regard: 'He's a hopeless schizophrenic,' for instance. Or these considerations show that, in the circumstances, the demand was not in effect: 'He has been under very great strain lately' (p. 78). For ease of reference, call this second way of inhibiting resentment an *exemption*.<sup>14</sup>

I want to emphasize that this taxonomy follows from Strawson's analysis of resentment and the like as, most fundamentally, reactions to a *demand's* being *violated*. A reaction to a violation of a demand may be inhibited if the demand was not in fact violated or if the demand itself is undermined.<sup>15</sup>

Strawson argues that neither sort of consideration operates universally. It is, first of all, simply isn't credible that a consideration of the first sort would do so: it simply isn't credible that everyone always satisfies the demand for goodwill.<sup>16</sup>

The more pressing question is whether an *exemption* might operate universally: whether the demand for goodwill, and hence all reactive attitudes (including not only resentment and indignation but also positive reactive attitudes, like gratitude, that we might have when the demand is satisfied) might be lifted altogether.

Recall that demanding goodwill of someone is an essential part of being in an interpersonal relationship with her. If I cease to demand goodwill of someone, to that extent I withdraw from my interpersonal relationship with her. I view her objectively, as an object to be manipulated or controlled rather than as a person to be engaged with. Exempting everyone from the demand for goodwill all the time, then, entails ceasing to participate in any engaged interpersonal relationships whatsoever.

Strawson argues that an across-the-board withdrawal from interpersonal relationships on these grounds is not possible for us. But just why this is impossible is not straightforwardly clear; Strawson seems to make several different moves at this point. Because the main argument of this article will be consistent with any reading of Strawson's claims on this point, I will only sketch briefly the several distinct claims that Strawson appears to make.

First, appealing to 'the human commitment to participation in ordinary inter-personal relationships,' which is 'thoroughgoing and deeply rooted' (p. 81), Strawson seems to say that, as a matter of natural fact, humans are incapable of withdrawing from all our interpersonal relationships. Relatedly, speaking of 'the human isolation' (p. 81) which this would entail, and appealing to 'an assessment of the gains and losses to human life' (p. 83), Strawson also seems to say that doing so would have unacceptable consequences. Finally, some of Strawson's remarks also suggest that, as a *conceptual* matter, we cannot imagine withdrawing from these relationships altogether on the grounds of a consideration like the truth of determinism. Interpersonal relationships (and the phenomena essentially connected to them) are, Strawson says, 'part of the general framework of human life, not something that can come up for review as particular cases can come up for review within this general framework' (p. 83).<sup>17</sup> The existence of interpersonal relationships is not subject to revision in light of reasons, because it is part of the framework within which such reason-driven revision makes sense. Rationally undermining the framework which gives sense to rational appraisal is, in some sense, incoherent.

For one (or more) of these reasons, it is not possible for us to withdraw from all interpersonal relationships. (Or, on the last option, it is not possible for us to have reasons for withdrawing from all such relationships.) Because participation in these relationships entails demanding goodwill, it is similarly impossible for us to drop all such demands. And finally, since that universal exemption is the only way that the reactive attitudes like resentment might be eliminated altogether, it is not possible to eliminate those reactive attitudes altogether.<sup>18</sup>

This argument relies on there being very tight connections between interpersonal relationships, the demands we make within such relationships, and our susceptibility to reactive attitudes like resentment: participation in interpersonal relationships requires demanding goodwill, and demanding goodwill requires that we be susceptible to reactive attitudes like resentment when the demand is violated. Those connections enable Strawson to extend his claim about the status of interpersonal relationships to the reactive attitudes.<sup>19</sup>

I want to focus on the connections between these three phenomena. If Strawson is correct that they are this closely linked, the Gandhi-King ideal is not, after all, so ideal. Eliminating our susceptibility to resentment

altogether, because it seems dangerous or morally problematic in some way – say, in response to the considerations raised by Gandhi and King – would require an impossible and unattractive exit from interpersonal relationships.

Additionally, if Strawson is correct that these three phenomena all stand or fall together, eliminating the susceptibility to resentment in any individual situation will carry the high cost of exiting a specific interpersonal relationship. If I am no longer susceptible to resentment in my interactions with someone, I do not demand goodwill of her, and I thereby adopt the objective attitude towards her. Doing this in a particular situation is not impossible in the way that adopting the objective attitude universally is impossible – we can and do adopt the objective attitude in a variety of situations<sup>20</sup> – but such a retreat to objectivity has its costs.

## 2.

To evaluate these claims, we need a better understanding of resentment's connection to the practice of making demands within interpersonal relationships. In this section I develop a detailed account of the roles that specifically antagonistic attitudes, like resentment, actually do play in our participation in engaged, demanding interpersonal relationships. This advance in our understanding of the significance of antagonistic response will enable me to describe a nonantagonistic alternative to resentment that can play the very same roles.

When Strawson introduces the phenomenon which he will go on to label the demand for goodwill, he describes it as a *desire*, pointing out '*how much we actually mind, how much it matters to us, whether the actions of other people . . . reflect attitudes towards us of goodwill, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other*' (p. 76, emphasis added).

Antagonistic response to the absence of goodwill does express this desire. But of course there is much more to demanding something than simply desiring it. To see what more is involved in making a demand, consider some specific scenarios in which one has desires about someone's attitudes but fails to demand anything of her. These are scenarios in which one takes the objective attitude towards someone – and so fails to engage in a demanding interpersonal relationship with her – even though one is concerned about her attitudes.

A specific example will be helpful. Consider a racist. Suppose that this person was raised in an intolerant social setting – say, in an economically depressed coal-mining town in eastern Kentucky. In virtue of the social and economic pressures that this person experienced growing up, she

might be unacceptably intolerant towards my ethnic group: she might unjustifiably blame us for unemployment and other social ills.

Now suppose that I respond to the racist with pity and charity. *You poor thing, I say, it's terrible, the pressures you were under growing up. Look, I don't take offense, because I know how your background led you to where you are now. But you're living a very impoverished life. You're missing out on valuable friendships; you're stuck in an unrealistic and angry way of viewing the world. You'll be much better off if I help you to improve your outlook on life.*

Even if this benevolent concern for the wrongdoer's wellbeing is sincere, it might strike us as a condescending or belittling way of responding to her. (This combination of benevolence and condescension may explain why it is attractive to take this stance when comforting a decent person who has been insulted or mistreated by a racist. One is still being *nice* to the racist, even while one is dismissing her and her attitudes as unimportant.) I am benevolently concerned about the racist's attitudes, and about her wellbeing; but there is something intuitively condescending about the way in which I am concerned. This is an indication that my concern – no matter how sincere – is what Strawson would call an objective, detached, disengaged way of relating to the wrongdoer.

Talk of condescension and belittling may seem to indicate that there is something at least *prima facie* morally problematic about taking this stance towards someone. I set that issue aside. My task here is to identify the specific reason that this particular way of relating to the racist is objective rather than engaged. The condescending quality of my attitudes toward the racist is significant only as an indication that my concern for her is objective.

What is missing in this case is *vulnerability*. I am only concerned about the racist's attitudes for her sake; her ill will toward me is not a threat to my wellbeing. I do not see the racist as a potential threat; I only see her, as Strawson puts it, as an 'object of social policy' to be 'cured' (p. 79).

An absence of vulnerability seems to explain some cases of disengaged concern about another's attitudes.<sup>21</sup> Being vulnerable to someone's attitudes contributes to participating in an engaged, demanding interpersonal relationship with her.

This is not, however, the only way of taking a detached, objective stance toward someone. Suppose that I do in fact find it quite unpleasant to be the target of the racist's ill will. *How awful, I might say. This person is really quite painful to be around. Her mind has been systematically perverted by the pressures she experienced growing up: her view of the world is so warped that she has lost touch with reality. She needs treatment – and in the meantime, I'm going to avoid her as much as I can.*

This does involve vulnerability to the racist's attitudes. But this is still a disengaged, objective way of relating to her. I view her as merely a

threatening or dangerous part of my environment: ‘something certainly to be taken account, perhaps precautionary account, of’ (p. 79), as Strawson puts it.

I am relating to the wrongdoer objectively in this case because I think she is unhealthy and warped.<sup>22</sup> Because I view her as impaired in this way, I exempt her from the demand for goodwill.<sup>23</sup> This indicates a second way of falling short of demanding goodwill of someone. Even though one cares about a person’s attitudes and is vulnerable to those attitudes, one may fail to view her as a healthy, full-fledged agent.

By design, the formulation of this condition is left somewhat schematic. Exactly what is required for agency, and exactly what is involved in falling short of being a healthy, full-fledged agent, are difficult substantive questions. I intend to formulate this condition in such a way as to be compatible with many ways of answering that question.

Demanding goodwill of someone within an engaged interpersonal relationship requires being vulnerable to the quality of that person’s will. It also requires viewing her as a healthy, full-fledged agent. But even when satisfying these requirements another form of disengagement is possible. Suppose that I am vulnerable to the racist’s ill will. Suppose further that I do take the racist to be a full-fledged, healthy agent. I do view her as an agent whose ill will may injure me. Even so, my response to her ill will might be quite *manipulative*. Perhaps I notice that the racist has a special aversion to direct confrontation. I might then be intentionally confrontational with her, with an eye to shaping and controlling her behavior toward me.

When I treat the racist in this way, the conditions identified above are satisfied: I do view her as a healthy, full-fledged agent, and I am vulnerable to her. But I am once again relating to her in a disengaged, objective way. Although I recognize her agency, I treat it in an objective fashion. It is condescending of me to attempt to shape the racist’s behavior by manipulating her, rather than engaging with her as a person, demanding that she act well, and then allowing her to respond. I am treating her agency as something to be ‘managed or handled . . . or trained’ (p. 79), or ‘something to be understood and controlled in the most desirable fashion’ (p. 82).

Again, I am not interested here in whether we would be justified in treating the wrongdoer in this objective, condescending way. I only want to understand why this way of relating to the wrongdoer is objective. That is because I am violating a third requirement on demanding goodwill and participating in an interpersonal relationship with someone. In addition to viewing someone as a healthy, full-fledged agent, one must *treat* her as such an agent. That requires that I avoid taking a manipulative attitude towards the racist.

I began this section with an observation: when the demand for goodwill is violated, angry, antagonistic reactions to that violation, like resentment,

seem to do important work keeping that demand in place (ensuring that it is not dropped or weakened in the aftermath of wrongdoing) and maintaining an engaged interpersonal relationship with the wrongdoer. Now I have described three conditions on making a demand within an engaged interpersonal relationship. Those conditions enable us to see clearly the work that angry resentment performs. Angry resentment expresses vulnerability; it involves viewing the wrongdoer as a healthy, full-fledged agent; and it treats her as such an agent.

At this point one might worry that my characterization of demands made within interpersonal relationships is incomplete. One might think that something more is required by participation in an engaged, demanding interpersonal relationship than the vulnerability, viewing the other as a healthy, full-fledged agent, and treating her as such an agent (and so avoiding being manipulative) that I have identified. I cannot rule that out decisively. If there are further conditions on participation in interpersonal relationships, I would like to hear about them. Such conditions would add to our picture of the work done by angry, antagonistic reactions to ill will.

### 3.

Now I can state more precisely my central question: is an antagonistic reaction like anger required to do this work – to satisfy these conditions – in the aftermath of a violation of the demand? If there are nonantagonistic reactions capable of doing the very same work as antagonistic reactions, switching to such reactions would be a way of preserving interpersonal relationships while eliminating antagonistic reactions.

I think we can conceive of some such nonantagonistic alternatives. Consider yet another way of responding to the racist. In light of the racist's upbringing, I might modify my reactive attitudes. Without taking that upbringing to have impaired the racist's agency, my resentment of her might diminish or dissipate altogether. *It's sad how poorly you're treating me, I might say. I expect better of you, and you've fallen short. I'm not angry with you, but I am very disappointed.* While still demanding that the racist be tolerant and goodwilled, my resentment might be replaced by sadness or disappointment at the racist's ill will.

I will argue that this shift would not involve giving up the demand for goodwill or exiting an interpersonal relationship, because this sadness response can satisfy the three conditions I have identified as well as anger can. To begin, notice that it is not anger *simpliciter* that keeps us within demanding interpersonal relationships. One can be benevolently angry at the plight of an agent to whom one is not vulnerable, or angry at the dangerous behavior of a seriously impaired agent. It is *resentment* that expresses vulnerability, and it is *resentment* that involves viewing and

treating its object as a healthy, full-fledged agent. It is *resentment*, not anger in general, that satisfies those necessary conditions on participation in engaged, demanding interpersonal relationships.

Resentment is a subclass – what Justin D’Arms and Daniel Jacobson call a ‘sharpening’ – of the more basic emotion of anger.

The claim that resentment is an emotion raises broader questions, about the nature of emotions, that are D’Arms and Jacobson’s primary interest; to avoid being drawn into those questions, I set to one side the claim that resentment is an emotion. Instead I want to draw on their useful framework for understanding the relationship between broader phenomena with affective and cognitive elements, like anger, and narrower phenomena, like resentment.

D’Arms and Jacobson illustrate the sharpening phenomenon that they have in mind with ‘tenure-denial anger’ or ‘tenure rage,’ which is the anger one has that one has been denied tenure. This is a *cognitive* sharpening: the subclass of anger is constructed by specifying a belief that all members of the subclass happen to share.<sup>24</sup> There are other ways of creating sharpenings of emotions, D’Arms and Jacobson indicate, for instance by the emotions’ causes or by their motivations.<sup>25</sup>

D’Arms and Jacobson propose that resentment is a cognitive sharpening of anger, distinguished by the belief that one has been wronged.<sup>26</sup> Since I take seriously Strawson’s focus on the demand for goodwill and involved interpersonal relationships, I would prefer to distinguish resentment not by the belief that one has been wronged, but rather by the specific beliefs and causal connections that I have just identified as essential to participation in such relationships. As a rough initial proposal, this might be captured by a simple cognitive sharpening of anger: resentment is a subclass of anger, distinguished by the belief that one has been left worse off by the ill will of a full-fledged, healthy agent.

This way of specifying the sharpening characteristic of resentment enables us to explain why angry resentment satisfies the three conditions set out above. Consider once more the example of the racist who bears me ill will. Suppose I react to the racist’s ill will with a sharpening of anger that includes this cognitive content. This sharpening of anger does reflect my vulnerability to her bigoted attitudes: this sharpening is, after all, anger at having been made worse off by such attitudes. This sharpening also involves viewing the racist as a healthy, full-fledged agent: it is characterized by the belief that she is such an agent. And this sharpening of anger treats the wrongdoer as a healthy, full-fledged agent. There is nothing manipulative in believing that someone’s attitudes have left one worse off, and being angry about that. (Of course, such an emotional reaction may be used to manipulate someone: if I know that the racist is so averse to hostility that she will do almost anything to avoid expressed anger, I may cultivate my resentful reaction to her behavior, in order to manipulate her

into avoiding such behavior in the future. But then it is my intention to cultivate my resentment with this aim in mind that is manipulative, not the resentment itself.)

I want to emphasize, however, that I am not committed to the details of this analysis of resentment. Perhaps resentment involves further cognitive content; or perhaps it is, in part, a causal sharpening of anger: perhaps resentment is anger which has a distinctive causal history. It is not necessary for my purposes to get these details of the sharpening that characterizes resentment exactly correct. The important general point is that *whatever sharpening distinguishes resentment from anger in general, we can apply the very same sharpening to sadness instead of anger.*

Sadness, like anger, is an affective response to something that one takes to be bad or unfortunate in some way.<sup>27</sup> We can describe a subclass of sadness that is set apart from sadness in general by the belief that a full-fledged, healthy agent has shown one ill will, which left one worse off. Or we could describe a sharpening that is distinguished by its causal origin in the attitudes of a full-fledged agent which leave one worse off. Or – to take D'Arms and Jacobson's account of resentment – we could describe a sharpening of sadness that is distinguished by the belief that one has been wronged.

This resentment-like sharpening of sadness might be called a form of disappointment (although I do not claim that it is the same as the natural phenomenon of disappointment).<sup>28</sup> Because, by stipulation, it involves the same elements that ensure that resentment keeps one within an engaged, demanding interpersonal relationship, this sharpening of sadness also keeps one within an engaged, demanding interpersonal relationship. If I respond to the racist's ill will with unsharpened sadness – or, for that matter, unsharpened anger – I am not responding to her as a participant in an engaged interpersonal relationship. I can, however, respond to her ill will with a sharpening of sadness that is related to sadness in general as resentment is related to anger in general. That response satisfies the conditions on participation in an interpersonal relationship, and so avoids being condescending or disengaged, in just the way that resentment does. When I am *disappointed* with the racist, in a way that parallels being *resentful* of her, I continue to demand goodwill of her and maintain an engaged interpersonal relationship with her.

Some space does remain to resist this conclusion, but it is tightly circumscribed. Earlier, to characterize the work done by antagonistic reactions, I identified three necessary conditions on participation in an engaged interpersonal relationship: vulnerability; viewing someone as a healthy, full-fledged agent; and treating her as such an agent. At that point I acknowledged that this list might be incomplete: there might be some other requirements on participation in such a relationship, which antagonistic reactions satisfy.

That alone might seem to enable resistance to my conclusion, but it does not. For if the proposed additional condition is satisfied, not by anger *simpliciter*, but by *the sharpening of anger that characterizes resentment*, then the parallel sharpening of sadness is equally well-equipped to satisfy that additional condition. What is needed is an additional condition that is satisfied by mere antagonistic anger, not by the sharpening that characterizes resentment. In addition, this condition must be such that it can only be satisfied by an antagonistic response. I do not see why interpersonal relationships would require this, although I do not have a conclusive argument against the possibility. So I conclude that, setting aside such possibilities, replacing resentment with the parallel sharpening of sadness I have identified would not involve giving up the demand for goodwill.

I take it to be clear that this possible sharpening of sadness would not be an antagonistic emotion in the way that anger is: it does not involve the antagonistic motivations or affects characteristic of anger. And yet it does the same work to keep us in demanding, engaged interpersonal relationships. Hence (again, unless there are further connections between engaged relationships and antagonism that I have overlooked) our participation in such relationships does not stand or fall with our susceptibility to antagonistic reactions. If we were to shift from resentment to this sharpening of sadness, and thereby eliminate our susceptibility to antagonistic reactions, we would not thereby drop the demand for goodwill or exit our interpersonal relationships.

#### 4.

I have described a conceivable shift in our reactions to wrongdoing, from antagonism to a sharpened form of sadness. If making this shift does not entail dropping the demand for goodwill or withdrawing from interpersonal relationships, then whatever considerations Strawson uses to rule out exiting from interpersonal relationships altogether, or dropping the demand altogether, fail to rule out dropping resentment altogether.

Although I have positioned my argument in opposition to Strawson, it is not always clear that he would resist my argument. He does pay extensive, recurring attention to the antagonistic reactive attitude of resentment. And yet near the beginning of his discussion of resentment, Strawson indicates that he is using it as a stand-in for a broader range of attitudes. 'Resentment and gratitude are . . . a usefully opposed pair,' he says; 'but, of course, there is a whole continuum of reactive attitude and feeling stretching on both sides of these and – the most comfortable area – in between them' (p. 77). The way Strawson frames his use of resentment here at least seems to open the way for it to be a stand-in for a wider range of attitudes, including ones that are not antagonistic. And again, at the end

of his discussion, Strawson indicates that one should be 'chary of claiming as essential features of the concept of morality in general, forms of these attitudes which may have a local and temporary prominence. No doubt to some extent my own descriptions of human attitudes have reflected local and temporary features of our own culture' (p. 93).

At a single point late in 'Freedom and Resentment,' however, Strawson seems to indicate that antagonistic reactions are an essential part of the demand for goodwill. I mentioned this discussion at the beginning of this article; now it is worth quoting and considering at length:<sup>29</sup>

[T]hese attitudes of disapprobation and indignation [the third-personal analogues of resentment] are precisely the correlates of the moral demand in the case where the demand is felt to be disregarded. The making of the demand *is* the proneness to such attitudes. . . . The partial withdrawal of goodwill which *these* attitudes entail, the modification *they* entail of the general demand that another should, if possible, be spared suffering, is, rather, the consequence of *continuing* to view him as a member of the moral community; only as one who has offended against its demands. So the preparedness to acquiesce in the infliction of suffering on the offender which is an essential part of punishment is all of a piece with this whole range of attitudes of which I have been speaking (p. 90, Strawson's emphasis).

Strawson first identifies the demand<sup>30</sup> with susceptibility to particular attitudes (here, the third-personal, vicarious analogues to resentment, namely indignation and disapprobation). He then claims that those attitudes entail a *partial withdrawal of goodwill*, and a *greater willingness to acquiesce in the infliction of suffering*. Violation of the demand for goodwill entails a withdrawal of goodwill and a willingness to see someone suffer. Making the demand entails susceptibility to antagonistic attitudes.

I do not wish to reject this passage in its entirety. There is an important claim being made here that is, contrary to initial appearances, compatible with my argument. Strawson *identifies* a demand with the susceptibility to certain attitudes. I think he is correct in this. Demands are partly constituted by the attitudes they give rise to; it is impossible to separate a demand from its characteristic attitudinal expressions.<sup>31</sup> I do not dispute this connection.

But identifying demands with the susceptibility to particular attitudes seems to undermine my argument. I claimed that we can eliminate our susceptibility to antagonistic attitudes, without eliminating the demand for goodwill. Yet now I am agreeing that the demand for goodwill, like any demand, is partly constituted by susceptibility to particular attitudes. Then eliminating susceptibility to those attitudes must involve eliminating the demand.

My conclusions survive if they are formulated more carefully. In a loose sense, which I adopted in previous sections, we may speak of 'the demand which is essential to interpersonal relationships.' I argued that this demand

requires that we respond to violations of it with either (sharpenings of) anger or (sharpenings of) sadness. In a strict sense, though, we should individuate demands by their characteristic expressions. Then I should distinguish between ‘the demand which issues in (sharpenings of) anger’ and ‘the demand which issues in (sharpenings of) sadness.’ Speaking carefully, my argument in the previous section established that the sadness-issuing demand is capable of doing the very same work as the anger-issuing demand to keep us in engaged interpersonal relationships.

What we cannot imagine eliminating altogether, then, is not susceptibility to *resentment*, or *antagonism*; it is susceptibility to some attitudes which keep in place an interpersonally involved demand for goodwill. That demand may take different forms, involving different reactions. There is a real sense in which the demand issuing in (sharpenings of) anger is distinct from the demand issuing in (sharpenings of) sadness. But either demand – antagonism-generating or not – will do the same work to keep us within involved interpersonal relationships.

## 5.

At the beginning of ‘Freedom and Resentment,’ Strawson agrees with his compatibilist ‘optimist’ character that ‘the facts as we know them supply an adequate basis for the concepts and practices which the [incompatibilist] pessimist feels to be imperilled by the possibility of determinism’s truth’ (p. 73). As I understand it, Strawson’s argument supplies an ‘adequate basis’ for those concepts and practices which cannot be suspended without exiting an interpersonal relationship. I agree with Strawson that making some involved, vulnerable demand for interpersonal regard is secure in this way.

But I do not think that antagonistic responses are similarly secure. I have described a sharpened form of sadness that is analogous to resentment in just those respects which are required by the practice of making a demand within an interpersonal relationship. If there is such a nonantagonistic alternative to resentment, then Strawson’s defense of engaged interpersonal relationships and the demand for goodwill does not extend to antagonistic responses.

I think that clearly acknowledging this limit has intriguing consequences. If I am right, Strawson’s argument does not rule out a deep and far-reaching revision within our moral practices. Consider a universal version of the shift to disappointment. Suppose that, inspired perhaps by Gandhi and King, we universally respond to wrongdoing and ill will with disappointed sadness rather than resentment or other antagonistic attitudes. In responding this way we do not stop demanding goodwill of each other; people’s attitudes still matter to each other; we do not withdraw

from genuine, demanding interpersonal relationships with each other. But the texture of those relationships is dramatically changed. They are far less antagonistic.

As Gary Watson has pointed out, Gandhi and King seem to realize this ideal, without giving up on demanding goodwill: they attempt to ‘*stand up* for themselves and others;’ they ‘*confront* their oppressors,’ while ‘*urging* and even *demanding* consideration for themselves and others.’ But ‘they manage, or come much closer than others to managing, to do such things without vindictiveness or malice.’<sup>32</sup>

Some might doubt that Gandhi and King really managed to eliminate antagonism altogether. Perhaps eliminating antagonistic response is not, psychologically speaking, a possibility for human beings.<sup>33</sup> It might still stand as an ideal for human relationships, one that we endeavor to realize as much as possible, and in light of which we appraise – or disavow – our actual responses to wrongdoing.

I have not defended Gandhi’s and King’s claims about the moral urgency of this ideal. I have only focused on arguing that eliminating antagonism, as they demand, is compatible with remaining in engaged interpersonal relationships.

That makes arguing for the Gandhi-King ideal easier, in two ways. First, because Strawson’s claim (that certain features of human life are impossible to eliminate) does not extend to antagonistic reactions, we cannot immediately rule out a *universal* elimination of antagonism in the way that we might immediately rule out a universal exit from interpersonal relationships. Second, because dropping antagonistic reactions is separated from exiting interpersonal relationships, giving up the antagonistic reactions in an *individual* case does not require dropping the demand in that case and exiting that interpersonal relationship. The cost of dropping the antagonistic reactions to any extent is much lower than it would be if Strawson’s argument did extend to antagonistic reactions.

This second point makes a significant difference to appraisal of the Gandhi-King ideal. If Strawson’s argument did extend to antagonistic reactions, eliminating antagonism towards an individual would always involve exiting an interpersonal relationship with her. At the very least, that would provide us with grounds for doubting that eliminating antagonism is as commendable as Gandhi and King seem to think it is.

Even if Strawson’s argument is limited in the way I have described, it should inform and circumscribe the Gandhi-King ideal. The modification of our reactive attitudes that King and Gandhi recommend might be momentous and far-reaching; but it cannot involve abandoning some of the most basic aspects of human life. It cannot, for instance, involve abandoning our vulnerability to each other, as we might if we were to respond to wrongdoing with pure benevolence.

Before closing, it is important to recognize that Strawson's argument is not the only consideration relevant to appraisal of the Gandhi-King ideal. There may be other powerful reasons for rejecting the Gandhi-King ideal. Strawson points to the importance of our relationship with a wrongdoer. We might focus instead on the *victim* of wrongdoing. Fulfilling the Gandhi-King ideal by replacing antagonism toward a wrongdoer with the sadness I have described might seem disrespectful of, or unresponsive to, the importance of the victim of wrongdoing. Such concerns call for further exploration.

Strawson's 'facts as we know them' may provide an adequate basis for our demanding interpersonal regard; but they do not thereby provide a basis for angry, hostile, antagonistic response when the demand is violated. Our current – reactively antagonistic – form of life is not similarly secure. And the Gandhi-King ideal condemning such reactions is deeply attractive. If antagonistic reactions should remain, they need moral justification. If they should be eliminated, or condemned, we need to see why they are flawed. We have work to do.<sup>34</sup>

Whitney Humanities Center  
Yale University

#### NOTES

<sup>1</sup> Available at [http://nobelprize.org/nobel\\_prizes/peace/laureates/1964/king-acceptance.html](http://nobelprize.org/nobel_prizes/peace/laureates/1964/king-acceptance.html).

<sup>2</sup> Unless indicated otherwise, all citations in the text are to the reprint of Strawson, P. F. (2003) 'Freedom and Resentment,' in G. Watson (ed.) *Free Will*, 2<sup>nd</sup> edn. Oxford: Oxford University Press, 2003, pp. 72–93. There are other defenders of antagonistic response to wrongdoing, most notably Jeffrie Murphy. See, for instance, Murphy, J. (2003). *Getting Even: Forgiveness and its Limits*. Oxford: Oxford University Press. See also Jean Hampton and Jeffrie Murphy's jointly authored book Hampton, J. and Murphy, J. (1990). *Forgiveness and Mercy*. Cambridge: Cambridge University Press; Miller, W. (2006). *Eye for an Eye*. Cambridge: Cambridge University Press (an account of the role historically played by vengeance); and French, P. (2001). *The Virtues of Vengeance*. Lawrence, KS: Kansas University Press.

Others who write in Strawson's shadow have expressed hesitation about Strawson's claims in defense of antagonistic response. See for instance Watson, G. (2008). 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme,' in M. McKenna and P. Russell (eds) *Free Will and Reactive Attitudes*. London: Ashgate Press, 2008, pp. 115–141, at p. 140; Sher, G. (2006). *In Praise of Blame*. Oxford: Oxford University Press, p. 138; and Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press, pp. 212–213ff (discussed in more detail below). Seth Shabo [Shabo, S. (2012). 'Where Love and Resentment Meet: Strawson's Intrapersonal Defense of Resentment,' *The Philosophical Review* 121, pp. 95–124], responds to Pereboom from a Strawsonian perspective.

<sup>3</sup> Even this brief gloss stakes out a contestable interpretive position on 'Freedom and Resentment.' In section 1 below I present my reading of Strawson in greater detail. I do not spend much time on interpretive issues in the main text (although see footnotes 15–19 for

closer textual analysis, including an argument in note 18 that my reading of Strawson is neutral on one of the most contentious aspects of Strawson's discussion).

It is important to point out that, insofar as my argument targets a particular defense of antagonism that I find in Strawson, alternative readings of Strawson or alternative Strawson-inspired defenses of antagonism would not be touched by my argument. (I nevertheless do believe that the argument I describe is Strawson's, and that it is an extremely important approach to defending antagonistic response.) Thanks to an anonymous reviewer for prompting me to acknowledge this clearly.

<sup>4</sup> See Strawson, 2003, pp. 79ff.

<sup>5</sup> Strawson, 2003, p. 90. There, Strawson is speaking of indignation and disapprobation, the third-personal analogues of resentment; but he makes clear that his claims extend to resentment as well.

<sup>6</sup> As I point out in Section 5, this leaves untouched the possibility that we have other reasons, having nothing to do with the wrongdoer, for thinking that antagonistic response is important. Such response might, for instance, be required by our respect for the *victim* of wrongdoing.

<sup>7</sup> Pereboom, 2001, p. 98. Pereboom discusses alternatives to resentment and indignation further on pp. 199–207. Also see Shabo, 2012, which was published as the present paper was being completed and responds to Pereboom from a Strawsonian perspective.

<sup>8</sup> Interestingly, after Pereboom attacks responsibility and describes an alternative to resentment, he turns to describe some advantages of relinquishing moral anger in favor of that alternative. In doing so, he cites some of the moral concerns about anger – e.g. its destructive effect on relationships, and its tendency to spin out of control – to which King and Gandhi also appeal. But Pereboom continues to rely on his arguments that we do not have free will to provide the primary motivation for the switch from moral anger to moral sadness. See Pereboom, 2001, pp. 207–212.

<sup>9</sup> See Pereboom, 2001, p. 97 and pp. 200–201.

<sup>10</sup> See D'Arms, J. and Jacobson, D. (2003). 'The Significance of Recalcitrant Emotion (or, Anti-quasijudgmentalism),' in A. Hatzimoyis (ed.) *Philosophy and the Emotions*, Royal Institute of Philosophy Supplement 52. Cambridge: Cambridge University Press, pp. 127–145.

<sup>11</sup> This argument is open to the same sort of complaint that I level at Pereboom: someone could claim that my account in Section 2 does not adequately capture the connection between resentment and the demands made within engaged interpersonal relationships. I acknowledge this possibility at the end of Section 3, but argue that the way I describe moral sadness leaves the possibility for this sort of response extremely circumscribed.

<sup>12</sup> The boundaries between the personal and impersonal attitudes are unclear; I will not try to work them out here. What is important for my purposes is that Strawson clearly believes his argument applies to both categories; thus I draw equally on his parallel discussions of the two in what follows.

<sup>13</sup> See, e.g., Strawson, 2003, p. 80.

<sup>14</sup> Following the use of the term 'exemption' in Watson (2008) and Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press. This term is not used by Strawson, but it has gained wide currency in the literature.

<sup>15</sup> (This note and notes 17–19 below attempt to strengthen my claims about 'Freedom and Resentment,' by working through interpretive questions about the text in greater detail. Those not interested in the details of Strawson's essay may ignore them.) Strawson's discussion of this second category of resentment-inhibitors (what I call 'exemptions') in fact proceeds (on pp. 78–79, and again on p. 81) without any mention of the demand for goodwill. Strawson instead characterizes exemptions in terms of the 'objective attitude

(or range of attitudes)' towards someone – which involves, as I have said, exiting an engaged interpersonal relationship. This can create the impression that Strawson's exemptions involve exempting people from *relationships*, instead of exempting people from *demands*.

Exemptions from the demand for goodwill do involve exiting an interpersonal relationship with the person exempted; but, as will become apparent in the course of my discussion, it is important to recognize that we exit the interpersonal relationship *in virtue of the fact that we cease demanding goodwill*. I think the most compelling evidence for giving the demand the main explanatory role here is the fact that, as I point out in the main text, doing so makes clear sense of Strawson's separation of inhibiting considerations into two kinds. (In addition, when Strawson repeats his argument in his section 5 in application to the 'third personal analogues of resentment' like indignation, he does repeatedly describe exemptions as suspensions of the demand for goodwill. See p. 86.)

<sup>16</sup> Strawson also points out that eliminating resentment for this reason would not amount to undermining responsibility altogether, in the way that concerns incompatibilists like his 'pessimist' character. This way of universally eliminating resentment would simply mean that everyone – perhaps quite responsibly, and praise-worthily – bears everyone else the appropriate level of goodwill.

<sup>17</sup> In the text, Strawson says that 'our natural human commitment to ordinary interpersonal *attitudes*' – not *relationships* – is 'a part of the general framework . . .' (p. 83, emphasis added). When this suggestion recurs near the end of 'Freedom and Resentment,' it is once again stated in terms of attitudes:

Inside the general structure or web of human *attitudes and feelings* of which I have been speaking, there is endless room for modification, redirection, criticism, and justification. But questions of justification are internal to the structure or relate to modifications internal to it. The existence of the general framework of *attitudes* itself is something we are given with the fact of human society. As a whole, it neither calls for, nor permits, an external 'rational' justification (Strawson, 2003, p. 91, emphasis added).

I read these as claims that the general system of interpersonal relationships, constituted by a wide variety of attitudes, is beyond rational criticism. That system may be described as a web of attitudes and feelings, as Strawson does; but I think it best to characterize the system in terms of relationships rather than attitudes. That is because speaking in terms of attitudes here is liable to mislead.

In light of Strawson's focus on the reactive attitudes, it is easy to take him to be claiming that what is beyond criticism are the specific *reactive* attitudes he identifies, including resentment. But Strawson avoids making such a claim, speaking instead of 'ordinary inter-personal attitudes in general' (pp. 82–3) and 'the general structure or web of human attitudes and feelings' (p. 91). If Strawson felt that he could claim, straightaway, that the reactive attitudes in particular are impossible to eliminate, or beyond justification, there would be no payoff to connecting those reactive attitudes to the demand for goodwill or to interpersonal relationships in general. I do not think that those connections are irrelevant to Strawson's argument; they seem, instead, quite central to it. As I read Strawson, then, those connections serve to situate the reactive attitudes within the general structure, about which he feels he can make the 'it is beyond rational criticism' claim.

<sup>18</sup> Depending on the way we understand Strawson's rejection of complete withdrawal from interpersonal relationships, this argumentative move will vary in important ways. I want to emphasize, for those concerned about these matters, that *my main argument in this article will be consistent with any reading of Strawson's claims on this point*. Some consideration of these two readings of Strawson will help to make that clear.

If Strawson's rejection is just a claim about what is motivationally feasible for human beings, his argument will run along these lines: if some consideration (e.g. the truth of determinism) undermines the demand for goodwill, it thereby leads us to withdraw from all interpersonal relationships; but, we are motivationally incapable of withdrawing from all interpersonal relationships; so, no consideration can motivate us to eliminate the demand for goodwill. This version of Strawson's argument leaves it open that some consideration *ought* to undermine the demand for goodwill, and we *ought* to withdraw from interpersonal relationships, although we are motivationally incapable of doing so.

Alternatively, I have said, Strawson might be making a claim about the role that these phenomena play in structuring human life: they are part of 'the general framework of human life' (p. 83), which is required for evaluative claims to make sense. A rational evaluation of the entire framework simply does not make sense. If some reason leads us to think that all demands for goodwill are unjustified, and so to think that all relationships are unjustified, we are engaged in just that senseless rational evaluation of the entire framework.

Some of Strawson's discussion is strongly suggestive of this second interpretation. He writes of 'the general structure or web of human attitudes and feelings,' saying that 'questions of justification are internal to the structure or relate to modifications internal to it. The existence of the general framework of attitudes itself is something we are given with the fact of human society. *As a whole, it neither calls for, nor permits, an external "rational" justification*' (p. 91, my emphasis).

In addition, on this second reading the possibility left open by the first sort of argument – that we ought to drop the demand for goodwill altogether, but we are motivationally incapable of doing so – is closed off. There is no way to assert that we ought to drop the demand altogether, because that would entail that the general framework of human life is somehow rationally undermined, which is incoherent.

The latter claim – that interpersonal relationships and the demand for goodwill are part of the general framework within which talk of justification makes sense, so that it is not coherent to ask if the framework itself is justified – may seem obscure to some philosophers. (For one, it may seem that we can at least ask if the general framework is itself internally coherent or not.)

Exactly how to understand this portion of Strawson's argument has been a major focus of philosophical discussion of 'Freedom and Resentment.' But it is not my focus. Strawson's argument has two main parts: a claim that we cannot exit interpersonal relationships altogether, and a claim that interpersonal relationships, the demand for goodwill, and the reactive attitudes are so closely tied together that they stand or fall as one. I am focused on the second claim; in particular, I will attempt to separate the antagonistic reactive attitudes from the demand and the relationship. Then we can challenge the antagonistic reactive attitudes, without challenging the demand or the interpersonal relationships. We need not address Strawson's argument about the status of interpersonal relationships – or even settle its specifics – to be able to mount that challenge.

<sup>19</sup> It may be helpful to compare my reading of Strawson with the one R. Jay Wallace develops in his influential book (Wallace, 1994). Wallace accepts that the reactive attitudes and a certain sort of demand are closely connected elements of what he calls 'holding someone to an expectation' (see Wallace, 1994, pp. 21–25). Wallace thus accepts the connection that I will reject, between making demands of someone and reacting in an antagonistic manner. (I will, however, accept a more limited version of this connection: see section 3, especially pp. 22–23.) Wallace certainly does think that the reactions in question are antagonistic: Wallace (1994, p. 93) makes it clear that his project of normatively assessing holding people to expectations is predicated on the antagonistic character of such reactions. Wallace also rejects connecting the reactive attitudes to interpersonal relationships (see

Wallace, 1994, pp. 25–33). He supports this rejection by pointing out – quite rightly – that simply defining the reactive attitudes as those attitudes we have within interpersonal relationships will be quite unsatisfying. Curiously, Wallace does not, so far as I can see, consider tying the reactive attitudes to interpersonal relationships by way of an expectation or demand. That is, he does not seem to consider the claim – which I take to be much more plausible, and quite central to ‘Freedom and Resentment’ – that it is the *demand* which is essential to participation in interpersonal relationships; and so he does not consider the more indirect way of connecting the reactive attitudes to interpersonal relationships by way of their connection to that demand, which I attribute to Strawson.

<sup>20</sup> For instance, we do so with ordinary, healthy people who are so obnoxious that it is too much of a strain to maintain engaged interpersonal relationships with them. See Strawson, 2003, pp. 79–80.

<sup>21</sup> This rules out a Socratic view of wrongdoing, on which the victim of wrongdoing is not seriously badly off because her soul remains just. If vulnerability is important to engaged interpersonal relationships, and we cannot entirely give up such relationships, this Socratic view is problematic: to the extent that it does not involve vulnerability to the attitudes of others, it grounds an impossibly disengaged, objective way of relating to each other. (Howard Bloch has suggested to me that Nietzsche would also reject the vulnerability condition: Nietzsche would not think that a demanding relation with another requires any vulnerability to her. Of course, that might be part of a broader rejection of the Strawsonian concern with engaged interpersonal relationships.)

<sup>22</sup> Whether or not this disengagement is *condescending* will depend on whether or not this is an accurate description of the racist. If it is accurate, it seems entirely appropriate. If it is inaccurate – if the racist is a healthy, full-fledged agent, but I take her to be impaired – my stance seems condescending. I think my description of the influence the racist’s upbringing has on her attitudes leaves it open whether she is, in fact, impaired.

<sup>23</sup> Thanks to Andrew Khoury for pointing out to me that this is a natural way of reacting to the racist, and one that amounts to exempting the racist from the demand.

<sup>24</sup> See D’Arms and Jacobson, 2003, p. 137.

<sup>25</sup> See D’Arms and Jacobson, 2003, p. 138 n. 22. They suggest that spite and vengefulness are motivational sharpenings, though they do not pursue that suggestion.

<sup>26</sup> D’Arms and Jacobson, 2003, p. 143.

<sup>27</sup> Sharpening happiness or some other emotion in the same way is ruled out here. An interpersonal masochist who is happy to be the object of ill will would be vulnerable to that ill will; she might very well view and treat her tormentors as healthy, full-fledged agents; but she would not thereby demand goodwill of them. Her happiness, even though it might be sharpened in the appropriate way, would not express any desire that her tormentors show her goodwill instead.

<sup>28</sup> Thanks to Jeff Helmreich for suggesting this natural label for the sadness I describe.

<sup>29</sup> One might think that Strawson’s talk of ‘our natural human commitment to ordinary interpersonal attitudes’ (2003, p. 83; see also p. 91) indicates that we have such a commitment to the antagonistic attitudes in particular. I think that is a bad misreading of Strawson’s claim, however, which is carefully directed at the *general* ‘framework,’ ‘structure,’ or ‘web’ of attitudes, not at particular attitudes. (See also note 15 above.)

<sup>30</sup> Here, within the context of Strawson’s discussion of the vicarious, third-person analogues of resentment, the demand in question is ‘the moral demand,’ rather than the demand for goodwill; and it is situated within ‘the moral community’ rather than interpersonal relationships. Strawson’s remarks on p. 84 indicate that this demand made by the moral community is the same as the demand for goodwill made within an interpersonal relation-

ship. And on p. 86 Strawson indicates that 'the moral community' is itself a generalization of demanding interpersonal relationships.

<sup>31</sup> This has been recognized by others inspired by Strawson, most prominently R. Jay Wallace (Wallace, 1994; see e.g. pp. 12 and 20–25).

<sup>32</sup> At the end of his discussion of Strawson's argument in Watson, 2008, p. 140.

<sup>33</sup> On one reading of Strawson, he holds that eliminating resentment altogether is psychologically impossible for us, because doing so would involve exiting interpersonal relationships altogether, and that is psychologically impossible for us. As I indicate in note 16 above, I take my argument to have force against this version of Strawson's argument as much as any other. Here I am referring to a different position: that it is psychologically impossible to eliminate antagonistic reactions to wrongdoing, for reasons having nothing to do with their connection to engaged interpersonal relationships. Perhaps, for instance, some contingent fact about our neurophysiology means that judgments about wrongdoing will always generate antagonistic attitudes.

<sup>34</sup> I am grateful to Zac Cogley, Tiffany Cvrkel, Stephen Darwall, Daniel Haggerty, Barbara Herman, Pamela Hieronymi, Andrew Khoury, Gavin Lawrence, Calvin Normore, and two anonymous referees for detailed and insightful feedback on this paper during its development. Earlier versions were presented to the Central States Philosophical Association, the Albritton Society, the Rocky Mountain Ethics Congress, the Eastern APA, and the UCLA Ethics Writing Workshop. I am indebted to those audiences for rich and wide-ranging discussions of this material. Work on this article was generously supported by a UCLA Dissertation Year Fellowship and by an Andrew W. Mellon Postdoctoral Fellowship. I am grateful for that support.