6

Respecting Each Other and Taking Responsibility for Our Biases

ELINOR MASON

1 Introduction

One puzzling feature of a world like ours, one riven with various sorts of prejudice and oppression, is that oppressive behavior is not always malicious. An individual who has simply absorbed the attitudes of his social world might be entirely well meaning, honestly doing his best. The old-fashioned sexist, who takes his own behavior to be appropriate and chivalrous, has no idea that he is actually patronizing and belittling the women around him. From his point of view, he is acting entirely appropriately, and furthermore, he is justified in that view, given the world in which he has been raised. In that case, there is a puzzle about how and why such a person might be blameworthy.

The standard view about blameworthiness is that, in order to be blameworthy, it is crucial that the agent has bad motives of some sort: what philosophers tend to call a bad will. But once we think about the nature of the agent's environment in cases like the old-fashioned sexist, it is not clear that we can locate the problematic aspects of the agent's behavior in the agent's will—rather, the problem comes from outside, from the nature of the society. It seems that such agents are nonculpably ignorant of the relevant moral truths. Thus the bad act does not reflect a poor quality of will. Yet, when people act in sexist and racist ways, through implicit bias, or more or less explicit bias, we are not inclined to take such agents to be beyond reproach. So the question I attempt to answer in this chapter is, in what way can we blame such agents? What is the basis, if any, for blaming them?

The answer I give is that agents are sometimes blameworthy (where I really mean that they are blame*worthy*, and not just that it is permissible to reproach them), even if they do not have any bad will. I argue that although the

paradigmatic account of blameworthiness is based on quality of will, we can and should be willing to allow that there are nonparadigmatic cases. I argue that the zone of responsibility can be extended to include acts that we are not fully in control of, and acts whose moral status we are nonculpably ignorant about at the time of acting. This extension of responsibility happens through a voluntary *taking* of responsibility. In what follows, I argue that there are certain conditions under which we should take responsibility, and that when we do so, we genuinely are responsible.

I start by arguing that it is indeed possible that some racist and sexist agents act without bad will, and so there is a category of problematic behavior that is not covered by traditional approaches to blameworthiness. I argue that Cheshire Calhoun's suggestion (1989), that we should make a distinction between blame and moral reproach, does not go far enough in the cases that I am interested in. Intuitively, in these cases, we need something closer to full-blown blame, despite the admitted lack of bad will. Finally, I present my own suggestion, which is that we should think of a sphere of responsibility that is not based on bad will, but rather, is based on agents' *taking responsibility*. I argue that we can understand the reasons for taking responsibility in terms of our relationships with others, whether those are personal relationships, or the relationships we have in virtue of a shared society.

2   Quality of Will

The central compatibilist idea, an idea we originally see in Hobbes and Hume, that gets refined by more recent philosophers such as Strawson (1962) and Frankfurt (1971), is that the crucial thing for moral responsibility is the *quality of will* behind the act.[1] The phrase 'quality of will' (due to Strawson 1962) is now a term of art in this field, but I will very briefly explain what I mean by it. It would be a mistake to think that 'will' in 'quality of will' picks out something sui generis or mysterious. Rather, the word is being used to refer to an everyday notion: the motivations that lead to an action. So saying that what matters for moral responsibility is quality of will is just saying that what matters is how an agent was motivated.[2]

One major disagreement in the contemporary literature on compatibilism is over whether nonculpable ignorance is always an excuse. Volitionists argue that if an agent does not know what she is doing when she acts, she cannot be held responsible—she does not have the right sort of quality of will. According to volitionists, there is a control condition on responsibility, such that an agent must be in control of what she is doing, where that includes knowing what she is doing under the relevant description. For a hard-line volitionist, nonculpably

acquired ignorance is always an excuse. So an agent who does not know that the money she donates to charity is actually going straight into the pockets of oligarchs cannot be held responsible for that, and the agent who does not know that he is insulting someone by calling her 'young lady' cannot be held responsible for that, and so on.[3]

By contrast, attributionists argue that quality of will can be good or bad *without* self-conscious knowledge of what is being done.[4] Attributionists argue that we can be blameworthy or praiseworthy as a result of our unconscious motivations. In Angela Smith's example (2005), the agent forgets her friend's birthday. She didn't forget deliberately, and there is (let's assume) nothing she deliberately did that caused her to forget. But her forgetting is caused by her bad will, by the fact that she does not care enough about her friend. An agent may tell herself that she is acting well, but in fact be driven by bad motivations. Words of "kindness" are sometimes passive aggressive. Omissions and forgettings often indicate carelessness. A slip is sometimes Freudian. Thus, according to attributionists, we can say that the agent is blameworthy, even though she acted in ignorance.

My own view is that these are both ways of having a relevant quality of will, and although they give rise to slightly different sorts of blameworthiness, they are both important ways of being responsible.[5] One way to have a bad will is to act badly knowing exactly what you are doing. Another way to have a bad will is to have bad motivations that one is not aware of. Having bad motivations is a way of having a bad will, even when there is no history of self-aware, culpable acts that led to these bad motivations. Of course, it may not be easy to determine what is going on in a particular case. Even in one's own case it is not always easy to tell how deep motivations are working, and what knowledge one has at a subconscious level. But there is nonetheless a fairly clear foundation for blameworthiness here: blameworthiness depends on bad will, understood in one of these two ways.[6]

3   Bad Actions without Bad Will

In this section, I discuss three categories of bad action that (arguably) do not emanate from a bad will. In all three cases, the standard view is that we are not blameworthy for these actions. But we rightly, I argue, feel some unease with the standard view. Blameworthiness in these cases may be complex, but it is not obvious that there is no blameworthiness.

Some of the ways in which an agent can produce a bad outcome without a bad will are entirely innocent and uncontroversial—cases where the bad outcome is entirely down to obscurity or glitches in the outside world, and has nothing to do with agency per se. For example, an agent whose car breaks down

in a completely unpredictable way is not blameworthy for consequently missing a meeting. An agent is not to blame for wasting a charitable donation when the money is, through hidden and obscure means, going astray. An agent whose well-meaning words of kindness hit a raw nerve and offend someone in a way that could not reasonably have been expected is not blameworthy for the offense.

But there are also cases where although it is plausible that an agent does not have a bad will, it is not completely clear that the agent is off the hook. These are the cases that I focus on here.

3.1 Implicit Bias

One such case is the phenomenon of implicit or unconscious bias. I want to remain neutral on what sort of attitude or attitudes implicit bias is composed of.[7] All I want to commit to is that implicit biases are revealed in our behavior (our hiring behavior, for example) and that we are not aware of having these biases.

There is broad agreement, both in the volitionist camp and in the attributionist camp, that we can be *indirectly* responsible for our biases in a fairly straightforward way.[8] If we are aware that we have biases, or are likely to, and if we know that there are things we can do to avoid having or manifesting biases, then obviously we should do those things, and if we do not, we are blameworthy. This is just analogous to indirect responsibility for our car malfunctioning. If we know that the brakes need to be checked, and we don't check them, we are blameworthy for the accident that is caused by the brakes failing.

However, there is disagreement over whether we can be *directly* responsible for our biases, or for the actions caused by our biases, and this is a hard question even for attributionists. Our biases are usually the result of absorbing problematic ideas from our culture. Thus they are importantly different from character traits, though they may appear to work in the same way. An innately misanthropic person is like that by nature, and so although we do not blame him for acquiring the trait (after all, he did not acquire it deliberately or though carelessness), we do blame him for the behavior that manifests the trait. We also judge the person for *being* misanthropic—for having (though not for acquiring) the trait.[9]

An implicit bias is not usually acquired deliberately or through carelessness, and so the agent is not usually blameworthy for acquiring the bias.[10] However, the mere fact that a bias is caused by external forces does not mean that it does not implicate the agent's will. We must be careful not to confuse the origin of a trait with its nature. A misanthrope may be misanthropic because of a virus contracted in infancy, but that does not make a difference to how misanthropic he is. His quality of will is not affected by the fact of a certain history (and though of course different histories tend to produce different qualities of will, they may not). If we consistently stick to the claim that it is just the quality of will itself

that merits blame, then the *origin* of the trait cannot deliver an excuse.[11] The question, then, is not just where the trait comes from, but to what extent is the agent's will engaged in the exercise of the trait?

There is good reason to think that implicit biases often implicate the agent's will. People who test positive for implicit racist biases also tend to exhibit avoidance behavior, suggesting that there is something visceral going on, something that it would be hard to characterize without referring to the agent's deep motivations.[12] Compare a tendency to underestimate certain sorts of probability— the likelihood that two people in a room have the same birthday, for example. This tendency, which many of us have, is probably innate, but it does not seem to implicate our will in any interesting way. We do not make the error out of any suspect motivation. We just make it because our brains are pretty bad at dealing with probabilities. Implicit bias of the sort I am discussing here, by contrast, is very much involved in the moral texture of the world. It is very plausible that our deep motivations make us susceptible to implicit biases, that we are predisposed to accept hierarchies that favor us or that resonate with other morally suspect motivations.

So the attributionist account gets a foothold here. Our biases are very likely to go hand in hand with problematic motivations such as contempt, disgust, and so on. Thus there is a way to defend the view that implicit bias is, in itself, directly blameworthy: it may be a way of having a bad will. I think it is likely that many cases of implicit bias are like this.

However, there *may* be biases that do not involve bad will. Returning to the point that implicit biases seem to come from outside us, there is another way to understand this so that it makes more sense as a potential excuse. It might be that the operation of biases *bypasses* our will. We might just absorb information about stereotypes and reproduce it in our behavior without any attitude being involved at all. On this picture, we are, as animals, as machines, partly automated. I have no idea if that is correct, but it seems at least passingly plausible.[13] If our biases operate by bypassing our wills, then it cannot be said that we have a bad will in exhibiting bias.

In that case, neither volitionism nor attributionism could make sense of direct responsibility for biases. We would be left with a picture whereby, so long as we have taken all the measures that we reasonably could to avoid bias and its results, we would not be responsible for remaining biased. In the actual world, where we do not have good techniques for avoiding or ameliorating bias, that would leave a lot of bias and manifestation of bias blameless.

Perhaps this conclusion is acceptable, though it strikes me as unsatisfactory. Agency is complex. There isn't a clear dividing line between what we do through our will in the sense relevant for blameworthiness, and what just happens. Even if implicit biases are automated and do not engage my motivations, and so do

not reflect bad will, they nonetheless seem to issue from my agency. They do not seem to be in the category of mere events—if I unknowingly discriminate against the women on the shortlist, this is something *I have done*. It seems too quick to say that I am not responsible in any way. So we have a puzzle here—on the one hand, it might be that implicit bias does not involve bad will. On the other hand, there is a strong pull to find it blameworthy.

3.2 Explicit Bias

Explicit bias may not involve awareness of wrongdoing, but it also probably involves bad will most of the time, and so the attributionist sort of blameworthiness is applicable.[14] Both racism and homophobia seem intimately related to obviously problematic attitudes, like disgust and contempt. Racist and homophobic attitudes may be couched in terms of factual claims (about intelligence, or God's will, or whatever), but it does not seem likely that these factual mistakes are mere factual mistakes. It is fairly obvious that these mistakes are motivated by hostile attitudes.

Sexist attitudes, like racist and homophobic ones, also usually seem to depend on false factual beliefs. But it is very often plausible that sexist attitudes are not hostile and that they are not underlaid with contempt or disgust. Consider this example from Cheshire Calhoun:

> Imagine, for example, a man who always refers to women as "girls" or "ladies." He, too, is uncoerced into doing so and is in complete possession of normal adult reasoning faculties. Yet it seems he ought not to be blamed for linguistically infantilizing or patronizing women, for, from his point of view, one cannot reasonably expect him to see anything wrong with his actions. We may suppose that in his childhood, his father and mother referred to women as "girls" or "ladies." He may also have come to understand that the former is flattering because it suggests youth and the latter simply polite. We may suppose that the people to whom he was exposed when he was growing up gave him examples only of this linguistic use and this understanding of its significance. From his point of view, it is natural to conclude that "girl" is flattering rather than infantilizing and that "lady" is polite rather than patronizing. (Calhoun 1989, 398)

In this example, arguably, the man has no bad will at all. He is ignorant, for sure, but he is ignorant in an understandable way—he is not culpably ignorant. Nor is he exhibiting bad will in the attributionist sense. All his motivations toward women are benign. They are patronizing, and he would probably assent to

various claims that are patently false, such as that women are less good drivers than men. But from his point of view, his attitudes are apt, just as we believe (I hope correctly) that patting a dog, calling it a "good doggy," and praising it for sitting on command, are apt rather than patronizing. We can make a distinction between sexism and misogyny: not all sexism involves misogyny.[15]

There is an important point worth emphasizing here, and that is that the more historically isolated the case, the more likely we are to think that it is possible for the agent to be sexist without a bad will. Miranda Fricker (2007) considers an example of sexism that involves no bad will (though she does not frame it quite in those terms). In Fricker's example, taken from Patricia Highsmith's novel, *The Talented Mr Ripley*, the character of Herbert Greenleaf repeatedly discounts the views of Marge, the young woman who would have become his daughter-in-law had his son not been murdered. Marge is actually correct about the identity of the murderer, but Greenleaf repeatedly ignores her insights. From his point of view, Marge is not worth taking seriously (Fricker 2007, 14–15).[16]

Fricker's main argument is that one of the ways that a social hierarchy affects us is by affecting the credibility levels we accord to others.[17] As Fricker points out, and as others have argued before her, we do not give as much weight to the testimony of people who are lower down in the social hierarchy.[18] So the question arises, are people blameworthy for committing this testimonial injustice? Fricker argues that the agent is not blameworthy (100–101).[19]

Fricker's background thought is that Greenleaf's ignorance is non-culpable. His attitude toward Marge is not hostile or contemptuous on a personal level. He believes, and is justified in believing, that her perspective is inferior. We can imagine for the sake of argument (especially as he is fictional) that he has made no epistemic slips, there is nothing he missed, nothing going on that he is not in control of. The problem is in the world, not in him. Let's also stipulate that there is no bad will. It seems plausible that agents in solidly sexist eras are in the grip of cultural norms that are so powerful that the agent can accept them without any extra push from suspect motivations.

Things look different when the relevant evidence is readily available. Most of us now have evidence that we are all prone to biases, and that there is systematic discrimination against women and people of color, and so on. When the evidence is "available," it becomes less likely that someone who gets things wrong has no bad will. However, there are various reasons why available evidence might not be availed of. One reason is bad will. But there are other reasons: the agent might be isolated from the evidence by peculiar circumstances or by non-sinister character traits. In what follows, I will be working on the assumption that it is possible that even in our era, people can be in the grip of sexist conceptual frameworks that, though sexist, do not reflect a bad will.

I agree with Fricker that when the evidence is not available, there is some undermining of blameworthiness for false moral views.[20] Of course, the issue of evidence may not be necessary to establish blameworthiness: it may be that the attitude itself involves a bad will, and as I pointed out earlier, the attributionist account is sufficient to establish blameworthiness. But let's assume that it is possible to have sexist (or otherwise problematic) attitudes without having a bad will. My hunch is that in both the historical and present-day cases, we think that the agent is blameworthy in some way for those attitudes, even when the agent does not have a bad will. My aim in this chapter is to vindicate that thought. I also explain why an agent in the historical case is less blameworthy than the agent in the present day, given that (by hypothesis) neither has a bad will.

3.3 Glitches

I want to briefly mention one more area where there are bad acts without bad will. This category of acts is not connected with our prejudices, but with the flawed nature of our agency. Sometimes, we just forget things, or don't notice them, in ways that do not reflect a bad will. Randolph Clarke (2014) imagines a man who has promised his spouse that he will get milk on the way home. Imagine that there is nothing that he has failed to do that he should have done in order to remember. He could have set reminders, but it would have been "borderline compulsive" behavior to do so. He could have stopped himself from thinking about his work on the way home, but he often thinks about his work, and with no bad effects (2014, 165).[21] In other words, we are to imagine a case where an agent forgets something and there is no bad will. Forgetting the milk is not very important, but we sometimes forget things that are much more important, and again, it seems possible that there is no bad will in these cases. From now on I will call these unmotivated forgettings "glitches."

Many philosophers have argued that in this sort of case, if there really is no bad will, then the agent is not responsible.[22] Clarke disagrees; he says that the agent is responsible, and responsible just because he has fallen below a standard (2014, 167). The problem with that response is that it is not clear what 'falling below a standard' means. If it means that the agent did it knowingly, or with unconscious bad will, then we are back to the standard accounts of what renders an agent blameworthy: some sort of bad will. If, on the other hand, the agent fell below a standard without realizing it, without intending it, or without being poorly motivated with respect to that standard, then we have no reason to think the agent is blameworthy. There are many ways of falling below a standard that have no implications at all for blameworthiness—I might simply be too short to go on the rollercoaster.

In all of the cases I have mentioned (some implicit biases, some explicit biases, and glitches), we are pulled in two directions. On the one hand, the absence of a

bad will seems to indicate an excuse, as it usually does. In these cases, the will is not engaged in the action in the right way for the standard accounts of responsibility to get a hold. In some (but not all) cases of implicit bias (at least possibly), the will is bypassed by some heuristic the brain has. In some cases of explicit bias (again, the claim is not that all cases are like this, just that it is a possibility), so long as the beliefs are not inherently contemptuous, it is possible to accept them "innocently." In the case of glitches, obviously, there is just a glitch—sometimes we just slip for no reason.

On the other hand, it is not obvious that the agent is off the hook. Think about how we would feel on realizing that we have acted badly in one of these ways. Imagine being presented with evidence that our selection of job candidates was racially biased. Or imagine being presented with video evidence that we did not listen to a suggestion or argument from a female colleague, and yet when the same point was presented by a man we took it seriously. And of course, the experience of realizing that we have forgotten something important is familiar. In all those cases, I think, we are likely to feel something akin to guilt.

Introspection is not a completely reliable test here. First, we sometimes take on more or less guilt than we should—I return to that point in section 6. Second, what we should immediately do when presented with the evidence as in the above cases, and what most of us would immediately do, is examine our own conscience for bad will. Did I in fact forget your birthday because I don't care about you? Do I subconsciously undervalue people of color, or women? We might feel guilty because we assume that we do harbor bad will, even when we cannot find it, and it is probably a good thing that we are inclined to self-examination in these cases, and inclined to find ourselves guilty until proven innocent. But this is not the sort of guilt (or blameworthiness) that I seek to vindicate here.

So let's stipulate that in these cases there is no bad will. Let's imagine, further, that we are presented with the explanation for our behavior. Imagine that the mechanisms of the brain, the way it absorbs and encodes information, the way it glitches, and so on, are all explained to us. Should we now feel that we can safely distance ourselves from our actions, that our actions are not our fault, but rather are things that happen to us, just as if we were blown by the wind? I think not. I think it would be *apt* to feel responsible in these cases, and I will defend that view in the rest of the chapter.

4 Reproach

Let's return to Calhoun's discussion of the well-meaning man, whose basic framework for understanding the world is sexist. Calhoun herself does not put the issue in terms of whether the man has a bad quality of will.[23] She tacitly takes

for granted that the man does not have a bad will, and that therefore he is not blameworthy. Calhoun's concern is whether we should nonetheless blame him. Calhoun's answer is that we cannot let him off the hook, because that amounts to condoning his behavior, and it impedes moral progress. Calhoun argues that although this man is not strictly blameworthy, we should *reproach* him.

I am entirely in agreement with Calhoun's argument for reproach in this case, but as Calhoun herself says, that is not the same as saying that he is blameworthy. It is worth stressing this distinction. Our reasons for reproaching, that is to say, the act of reproaching, can differ from what we think of as the paradigm reasons for holding responsible without becoming the wrong sorts of reasons. The paradigm reason for holding someone responsible is that they *are* responsible. On the compatibilist account, that means that they have a relevant quality of will. But there can be other valid reasons for reproaching people.[24] For example, we may reproach children for things we don't actually take them to be responsible for in order to teach them how to *be* responsible. The sense in which Calhoun suggests we should hold the benignly motivated sexist responsible is not quite like that, because the context is very different, as she points out (1989, 401). But the justifying reasons for holding responsible are similar, in that they are forward-looking; they relate to the way we would like things to be, as opposed to being an assessment of what an agent has done in the past.[25]

My focus here is not on holding responsible; it is on *being* responsible, and on blame*worthiness*. In the next section, I argue that there is a way to make sense of blameworthiness proper in cases like the old-fashioned sexist.

5   Taking Responsibility

Before I present my argument for taking responsibility for biases, I need to make clear that I am talking about responsibility in the sense of blameworthiness, and not about *liability*. Liability, legal or conventional, is just the duty to make amends in the event that a bad outcome occurs. Strict liability (as it is called in law) does not imply that the agent who is liable must have a bad will, or even that she herself performs the relevant act: one can be liable for what one's tenant, or children do, and for what the wind does, or what just happens. Liability does not imply blameworthiness.

I am not talking about liability here. It may well be that we should take on liability in some of the cases I am interested in, but I am suggesting something more ambitious. I am suggesting that we should take on *responsibility* for these actions. We should *own* them: we should feel about them as we feel about the actions that we do deliberately, or out of good or bad will that we were not aware of at the time. We should be willing to extend the realm of our own responsibility.

Or, to put it another way, we should be willing to limit the excuses that we take ourselves to be entitled to appeal to. In general, we can appeal to various sorts of luck, accident, inadvertence, and ignorance as an excuse, but the kind of inadvertence that is in play here will not count as an excuse.

I shall motivate my claim that we should sometimes take responsibility by using an example.[26] Imagine a conversation between Angela and her friend, after Angela has forgotten an important birthday. Imagine that Angela searches her own conscience and finds no flaw in her will at all. Let's say, for the sake of argument, that she is accurate in her introspection. It is not the case that she failed to care enough. Rather, her forgetting her friend's birthday was just a glitch. Imagine that she now says to her friend, "Hey, you have nothing to worry about, it's not a big deal, I just forgot." There is something lacking in this response. The friend might justifiably feel that an apology is owed even if there was no bad will: she might feel, and with good reason, it seems, that the lack of bad will is not entirely exonerating.

We can imagine further, that Angela accepts that she should apologize. She accepts liability (though we do not usually talk about liability in contexts where restitution consists only, or mainly, of an apology). Her friend might still feel something is amiss. The friend might ask whether Angela feels bad, and meet this response: "No, I don't feel bad. As I said, it was a glitch. I have nothing to feel bad about. I have apologized. What more do you want?" There is a natural thought that her friend might reasonably reply, "I want you to feel bad about it! I want you to feel guilty—to take responsibility. This is not about you. It's about me. I hoped for a card from you on my birthday, and you just forgot!"

There are, of course, complexities in this story. One possible line of objection points to the fact that in most cases where we fail in a duty, even if it is completely clearly not blameworthy, it would be cold-hearted not to feel *anything* in response to the failure, but the appropriate feelings are not necessarily guilt and remorse; they are more like regret, sympathy, and empathy. The problem with Angela in this story, we might say, is just that she seems so cold-hearted.

It is true that when we fail in a duty, even when it is not through bad will, it is usually fitting to feel sympathy and empathy and so on, and it is true of course that we don't tend to like people who shrug off those feelings as unimportant. But there is more than that going on with Angela. Angela insists that she does not feel bad. She takes a very strict approach to her own responsibility—she knows that she had no bad will, and for her, that is the end of the story. But her friend wants more; their relationship demands more.

You might think that the friend is just wrong about things, and that Angela's clear-sighted assessment of the situation is preferable. But I am not arguing about what the reasons or evidence supports. I am talking about what we want from our loved ones, and the attitudes that we take as evidence that they are committed to us in the right sort of way. The idea here is that it is not always appropriate to

insist on the paradigmatic criteria for blameworthiness. As an analogy: the paradigmatic criteria for love might preclude loving our dogs. But the clear truth is that we do love our dogs. So we must adjust our conception of what love is.

In the example above, I focused on a personal relationship and the way in which personal relationships can have internal demands that do not align with philosophical conceptions of blameworthiness. Next, let's turn to failures to do one's duty to other people, where the duty is not a duty arising from a personal relationship, but just a duty we have to our fellow people. Start with implicit bias. Imagine that we discover that we are one of the subjects in the well-known hiring bias experiment, in which CVs associated with female names, or traditionally black names, are judged as less worthy than identical CVs that appear to belong to white men (Steinpreis, Anders, and Ritzke 1999; Bertrand and Mullainthan 2004). Imagine that you are confronted with the evidence that you have done this. In many cases, this may involve bad will, of familiar or more complex kinds.[27] But let's stipulate again, that in the cases I am discussing, there was no bad will. The act is like a glitch, in that it just happens, though of course the causal mechanism is slightly different in a glitch.

Again, there are complexities here. Of course you should accept liability: you should be willing to do your best to rectify the situation. And you should feel emotions like regret and sympathy. But, I argue, as a member of a society in which there are women and people of color, and a history of oppression, you should be willing to take on extended responsibility for this sort of failing. You should feel, not just regret, but something akin to guilt or remorse. Taking on that responsibility is constitutive of respecting your fellow community members in the appropriate way. Taking on responsibility reflects the gravity and nature of the offense. In treating someone as less valuable because of race or sex, we are undermining their personhood. We go some way to restoring it by taking on extended responsibility for our error.

Again, I will try to motivate the claim by using an example. Take an example of explicit bias: imagine that we explain to Calhoun's character (call him Stan) that "young lady" is patronizing, and that he should not think of men and women as having different roles and properties. Let's assume, as before, that Stan had no bad will, and that he approaches the issue with a willingness to believe that he has made a mistake. He tries hard to understand that age is irrelevant, and that gender neutrality is best, but that if there is no gender-neutral way to address someone, "woman" is better than "lady." Imagine two attitudes he might take to his own future errors:

> *Stan Doing His Best*:
> This is all very complex. I'm not sure that I am going to get it right.
> If I get it wrong, though, it's just an innocent mistake. Don't blame

me. I really am trying. I am doing my best. You can't ask more of me than that.

Stan Responsible:

This is all very complex. I'm not sure that I am going to get it right. I will feel terrible if I get it wrong in future, now that I understand more about all this, and I apologize in advance. But of course I am absolutely going to try my best.

Let me stipulate that each really is willing to try their best, that each *believes* that 'young lady' is patronizing. Nonetheless, Stan Responsible has taken the seriousness of the issue to heart, in a way that Stan Doing His Best has not. The fact that he will feel terrible shows that he is thinking about the women he will encounter in the future, and the way that his attitude to them will affect them. He is thinking of himself as having a relationship with other members of the community, a relationship that requires respect. So even though he knows that he may accidentally get something wrong in the future, and he knows that doing so would not betray a bad will on his part, he *takes* responsibility—he does not allow himself to appeal to inadvertence as an excuse.

By contrast, Stan Doing His Best seems more concerned with himself. He seems to have examined his conscience, found it clear, and takes it that doing his best is enough. What more can he do? He can't do more, of course, and my argument is not that he should do more than his best—he should do his best. But he should also eschew the excuse that further slips would be entirely inadvertent. He should take on responsibility.

Think about the way in which we are all Stan. We are, most of us, in a situation where we know that we are likely to be in the grip of various problematic views and assumptions, ways of framing the world and social relations that we have inherited and have not yet had our consciousness raised about. What should we be more concerned with: our own consciences, or the people to whom we are being unfair in our ignorance? The issue of whether or not we have a bad will is not the important issue. The important issue is that we take seriously that we are wronging someone. Taking it seriously involves taking responsibility for it, feeling bad about it.

And think about how we feel as the injured party. What do we want from our fellow members of society? We want them to do their best to avoid being biased against us, obviously. And I am not denying that that is the main thing we want. But that doesn't mean that we don't also want them to take seriously the wrongs they do us. Imagine a conversation between a woman and her male head of department, where he apologizes for having failed to promote her a few years ago. He explains that he now sees that although he did not have any bad will at all, he was unfortunately biased in various ways, and that he will now rectify the situation.

Again, he may or may not take on responsibility as well as liability. Understanding that he has actually taken responsibility, that he feels remorse, makes a huge difference to this exchange and to how things are in the future. Liability is not enough. Being a decent fellow member of a department, being a decent fellow member of a society, requires taking on responsibility for inadvertent injustices. So, when we inadvertently fail in our duties to others, where those duties are duties of relationships, personal or impersonal, we should take on responsibility.

The 'should' here is normative, but it does not indicate a moral duty. This is usefully illustrated by contrasting my argument with an argument presented by David Enoch. Enoch (2012) argues that we sometimes have a moral duty to take responsibility in areas where agency is 'penumbral'. Enoch's account of the range of actions for which we might take on responsibility is much wider than mine: Enoch talks about responsibility for what one's child does, for an accident that kills a pedestrian, and for what one's country does.[28] Thus our concerns are slightly different. I am concerned with actions that are genuinely an agent's actions. This illustrates an important problem in Enoch's account. The area of "penumbral agency" is not really an area of agency at all. It is an area in which we might reasonably be held *liable*, but it would be stretching things to say that we might be genuinely responsible, or blameworthy in my sense. Enoch conflates taking on responsibility with taking on liability.

The difference between taking on liability and taking on responsibility is a difference in the psychology of the agent—the attitudes and feelings that they are taking on. When we feel responsible for an action, we must feel that it is our action, that we did it, that we own it. Feelings of remorse and so on are apt when the agent herself acts badly, but are not apt in regard to the actions of others. Thus we cannot take on responsibility for the actions of others. We can take on liability. Taking on liability is just willingness to perform acts of recompense in the event of a bad outcome. It does not involve ownership, or the attendant feelings of remorse and so on.

It might be objected that what one should feel when one has inadvertently acted badly is not remorse, but something more like what Bernard Williams calls "agent regret" (1981, 27; 1993, 69). Indeed, Miranda Fricker, in a recent article, makes exactly this claim (2016). Fricker imagines someone who, through no epistemic fault, is in the grip of sexist prejudices. Such a person might be innocent in having acquired such bad views (they have come from the epistemic environment), and also justified in thinking that she is free of such prejudices. Fricker compares that person to Oedipus, as discussed by Williams (1993). Fricker says, "Both do voluntary things the significance of which, through no fault of their own, they do not grasp; and in both cases their failure to grasp it is down to their circumstances or environment. They both suffer a kind of *environmental* bad luck" (2016, 45).

The notion of agent regret is much weaker than the notion of remorse. It is also much broader. Williams (1981, 1993) uses various examples to justify his idea that agent regret is an overlooked moral emotion. One example depends on outcome luck: he compares two lorry drivers who drive with the same level of care: one kills a child, and one does not. The example of Oedipus seems a bit different, as Fricker points out. Oedipus acts deliberately, but he doesn't know what he is doing; he is unaware of some important facts. And what would we say about Helen of Troy—should she feel agent regret? Did she launch a thousand ships? Of course not, but at the same time, we might think it odd if Helen feels absolutely nothing about the fact that the Trojan War started because of her face. To be fair to Helen, her face did nothing—it was perceived in certain ways by certain people, and *they* started the Trojan War, but still, it was Helen's face, and even if her actions were beyond reproach, it would be odd of her to feel nothing at all. So we might be drawn to the view that there are shades of agent regret, varying in both quantity and quality.

This leads to the very natural thought that remorse as well as agent regret comes in various shades. Remorse is paradigmatically appropriate when one has knowingly and deliberately done something bad, when one is blameworthy on the volitionist account, in other words. As I said earlier, attributionists argue that one can be blameworthy without self-awareness. One way to understand the debate between volitionists and attributionists is as a dispute over what a univocal notion of blameworthiness applies to. But as Gary Watson (1996, 2004) influentially pointed out, we could think in terms of different sorts of responsibility.[29] I think that one is blameworthy in a slightly different way when one acts without awareness, but it is still genuine blameworthiness, and the agent should feel a very close relative of remorse. It would be a shame to conflate the distinction between these two ways of being blameworthy, but that does not downgrade either of them.[30]

So, returning to the distinction between taking on liability and taking on responsibility, the above reflections on shades of remorse do not undermine the distinction, even if they make us think that the boundaries might be vague. Taking on responsibility is more personal, more emotional, and more weighty than taking on liability. There is also an important difference in the sorts of reason one might have for taking on liability or responsibility. I have argued that the reason for taking on responsibility in certain cases is that it is constitutive of respectful relationships. There are some contexts in which, when we inadvertently let someone down, we have to own that. We have to take it on, give it the same status as the acts that issue from our will. We do not, however, have a *moral* duty to take on responsibility. We could eschew the whole relationship. The point is rather that if we are committed to the relationship, whether personal or social, we must act in certain ways.

Enoch (2012), by contrast, is interested in diagnosing a moral fault. He argues that the fault in the lorry driver who does not take on anything after (nonnegligently) having killed a pedestrian is that he has violated his moral duty to take responsibility (2012, 110–111). But although the lorry driver has a moral duty to take on liability—at the very least to apologize, and to make amends in other small ways if possible—he has no duty to take on responsibility or blameworthiness. Nothing hangs on him taking on more than liability—it is not required by the relationship in play.

Enoch thinks that taking on responsibility is an act of will, and that justifies his view that it is something we sometimes have a duty to do it. I agree if we understand him as talking about liability: in nonlegal contexts, taking on liability can be an act of will. But taking on responsibility is more like a disposition. Some people are inclined to take on responsibility; they do not actively *do it* so much as discover that they have done it. Ideally, their taking on responsibility is responsive to their relationships in an appropriate way. To respect your partner is (assuming that this is what your partner wants) to automatically feel bad about having let him down, even when you let him down inadvertently. To respect your fellow people is to feel bad when you wrong them.[31] Of course, our dispositions to take on responsibility are not always perfectly responsive to the genuine and legitimate needs of our relationships. They can be distorted in various ways. I return to that point below.

First, I want to briefly return to historical cases. I have argued that we should take responsibility for the ways in which we inadvertently let people down. But that doesn't clearly apply to people who never realize that they have let people down. Fricker's Greenleaf is just a man of his time, and he is trapped in his time. It seems that it does not make sense to say of Greenleaf that he should take responsibility for his mistakes, because he is not even remotely aware of the possibility that he is making that sort of mistake.

I agree that it does not make sense for Greenleaf to take responsibility. Ideally he would come to see things differently, of course, but if we imagine him in the historical context he is in, he does not have enough distance from his own oppressive actions and tendencies to take responsibility for them. Thus, my view makes sense of our ambivalence about historical cases. First, we can understand why it might be that the agent has no bad will—the epistemic circumstances are bad. And second, whereas we expect, both normatively and predictively, our peers to take responsibility for their inadvertent oppressive actions or tendencies, we know that someone who is historically isolated will not have done so, and will have had very limited opportunity to do so. As Fricker remarks, it would take someone really exceptional to see through the orthodoxy of the day (2007, 101). This is not to say that we have no critical resources. We still think, of course, that the agent should not have been so biased, that he has done

something wrong in being so biased, and so on, and that the world itself is problematic and needs to be changed.

There is one last point to make, and that is that when an agent takes on responsibility, she *is* responsible. One might think that taking on responsibility is artificial in some way, or, like taking on liability, it does not really involve the deep sense of responsibility. But my argument is that one can change one's standing with regard to an inadvertent action, and if one takes on responsibility for it, one really is responsible. This is not a radical claim: it is just a development of what the compatibilist needs for the compatibilist view to get going. Compatibilism requires a certain amount of good faith. It is always possible to turn around and suddenly see things in the light of determinism. If we do that, we can stop seeing ourselves as responsible at all. So the argument I am making here is just a development of that point; it builds on what compatibilism must already rely on: that we are willing to see ourselves as responsible.

6 Feminist Implications

I shall close with some feminist implications of this proposal.[32] It is a "well-known fact," in the sense that is likely to be taken for granted in popular cultural forms such as magazine articles and sitcoms, that women feel more guilt than men do.[33] We feel guilty about our children, our spouses, our pets, about what we eat, about how we speak, about how we look, about what we say, about how hard we work, about how messy our houses are, and, even more disturbingly, it seems that we feel guilt if we are assaulted or raped or harassed. It is always our fault. Men, by contrast, at least in the popular imagination, breeze through life, feeling guilty only about what they can control.

The alleged problem (it may be a figment of the popular imagination, but I will assume that the impression is accurate for the sake of argument) is not simply that women feel worse about what they have done. It is that women take on responsibility for more of what happens than men do. And that is not simply a matter of taking on more responsibilities (though it seems that women do that too, as the well-documented phenomenon of the "second shift" illustrates[34]). The point is that, as I have argued, we have some leeway over what we take responsibility for, and women seem to take on more.

The social context of this picture is, of course, a patriarchal society. Thus, as I suggested above, the ways in which we are inclined to take on extended responsibility are themselves hostage to the social environment we are immersed in. If women really do feel more guilt then men, it is partly because women are blamed more, are *held* more responsible than men are. For example, efforts to prevent unwanted pregnancies, rape, and sexual harassment often focus on

educating girls. Girls are taught to think that it is up to them to learn how to say "no" effectively, how to insist on condom use, how to dress so as to avoid unwanted attention, and so on. So it is not surprising that failure to prevent bad things happening feels like a failure. Part of what is going on here is that women are unfairly being made to take responsibility, to feel guilty, about things that others do.

But my argument suggests another explanation for the (putative) fact that women feel more guilt than men. I have argued that we take on extended responsibility because our relationships require it. As I pointed out earlier, our close personal relationships require what we *think* they require: we make our own relationship conditions. We make them partly on an individual basis, but we also make them culturally. In some cultures, arranged marriages seem fine; in others, they do not. What we can accept and what will make us happy depend very much on what we have been brought up to expect. And what we will accept depends on our bargaining position. The bargaining positions of those in oppressed groups are, of course, weaker than those of the privileged.

So it might seem that my view leaves little room for normative criticism of relationships that are shaped by oppressive social structures. If what people expect from women in a relationship is more responsibility than what they expect from men, that *makes it true* that the relationship requires more from women. And this is indeed the lived experience of many women in heterosexual relationships. In order for the relationship to work, to feel right, the woman has to take on an awful lot of extended responsibility.

The solutions are, of course, not out of reach. I find it interesting that once we understand responsibility as involving an element of taking responsibility, and once we see that that is connected to our social relationships, it is evident that the nature of our social relationships must be examined carefully. We cannot simply take for granted how we respond to others; rather, we should interrogate our responses. We cannot take for granted the shape of our relationships, but should examine them for ideological distortions. My view is not that we should take everything for granted and base responsibility on that. My view is rather that *especially* because responsibility is intertwined with our social world, it is important that we are critical of our social world.

7 Conclusion

I have argued that there is an interesting category of acts that do not betray bad will, but that are nonetheless not clearly ones that we should let people off the hook for. Such acts are *our* acts, in that they come from our agency in a broad sense. Some forms of implicit and explicit bias fall into this category. I have

argued that we should *take* responsibility for these actions in the situations where doing so is necessary for the relevant relationship. So if there is no relationship at all, taking responsibility is not required. What we should take responsibility for varies with the relationship. In personal relationships, we should take responsibility for a wide range of glitches as well as our biases. In our broader social relationships, we need to take responsibility for our biases. Taking responsibility is constitutive of respecting in the appropriate way.[35]

Notes

1. I am not claiming that all compatibilists would accept this as a characterization of the view. My point here is that quality of will is the *central* feature of a compatibilist account. In fact, many compatibilist views depart from a quality of will view, bringing in notions such as history and capacity that are not intrinsic qualities of an agent's will, but extrinsic ones. There is also disagreement over what sort of quality of will is relevant, and not just over whether self awareness is necessary, but over what levels of reasons responsiveness are needed and how reasons responsiveness is understood, and over what sort of identification with one's own will is required.
2. See also Michael McKenna's very clear account of what 'quality of will' means (2012, 58–59), though I disagree with McKenna on whether 'quality' only refers to moral worth.
3. See Wallace (1994, 138–139); Zimmerman (1997 and elsewhere); Rosen (2002 and elsewhere); Levy (2011 and elsewhere).
4. See Watson (1996, 2004) where the term 'attributionism' originates. The view is developed (in various different ways) by others, including Nomy Arpaly (2003 and elsewhere); Angela Smith (2005 and elsewhere); T. M. Scanlon (1998 and elsewhere); Matthew Talbert (2008 and elsewhere).
5. See Mason 2015.
6. I focus on blameworthiness and fault rather than praiseworthiness and achievement because in general we are more tolerant of luck in praiseworthiness; we are less reluctant to allow luck to play a role in praiseworthy outcomes because less is at stake for the actor.
7. See Michael Brownstein's *Stanford Encyclopedia* article (2016) for a review of the various views on that.
8. See, e.g., Kelly and Roeder 2008; Holroyd 2012; Vargas 2013, esp. chapter 8.
9. See Adams (1985) for a very clear justification of that outlook, which of course is the basis of the attributionist view.
10. Indirect responsibility for biases depends on the cultural situation—in a situation where there is lots of awareness of bias and good and well-known techniques for countering bias and avoiding acquiring bias, it will be much easier to assign blameworthiness for acquiring and having biases.
11. Gary Watson points out this problem (1987; 2004, 247).
12. Research by Fazio et al. (1995) shows a correlation between implicit racism and sitting further away from people of that race. Bessenof and Sherman (2000) demonstrate a similar correlation between bias against fat people and avoidance behavior. These and other results are discussed in Dasgupta (2004).
13. This view is suggested by the psychologist John Bargh (2005). I owe the reference to Holly Smith's discussion (2011, 135). Smith concedes that if there are indeed fully automatic processes, then we are not blameworthy for them, and her overall view is that we are blameworthy only if a sufficient proportion of our attitudes are involved in our behavior.
14. The volitionist has a harder time making sense of blameworthiness for explicit bias, because the agent who is genuinely in the grip of false views all the way down, who has never engaged in culpable self-deception, cannot be blameworthy on the volitionist view. This is one reason the volitionist view seems implausible when proposed as the sole account of responsibility.

15. And of course it is possible that we could tell a similar story about other biases, including racism and homophobia.

16. Fricker also draws on Anthony Minghella's screenplay, in which the character of Greenleaf says at one point, "Marge, there's female intuition, and then there are facts" (2007, 1).

17. I think that epistemic injustice can manifest as both implicit and explicit bias. It appears as implicit bias when the agent is not at all aware that she is listening less to, and taking less seriously, members of the oppressed group. It appears as explicit bias when people consciously believe things like "women are more intuitive than men and men are more rational then women."

18. See, e.g., Mills (1998), who argues the point in relation to race.

19. Fricker claims that to call the agent blameworthy in this sort of case would be hubris (2007, 105). Instead, she argues for something that we can justifiably do (or feel) in place of blame. She suggests that when someone's circumstances are epistemically limited, although we cannot aptly feel the resentment of blame, we can aptly feel "resentment of disappointment" (104). I return to this below.

20. Arpaly makes this point too: she gives an example of a young boy raised in an entirely sexist and segregated society (2003, 104).

21. Clarke is interested in responsibility for omissions, but for the sake of my argument here, nothing hangs on this being a case of omission. What I call a "glitch" is often called negligence, though that is misleading in various ways. For other examples with this structure, see Vargas 2005; King 2009; Sher 2009; Raz 2011; Tannenbaum 2015. See also Mason 2018, forthcoming.

22. Volitionists, of course, argue that the agent is not responsible in these cases, as the control or knowledge component is missing, and this case is, from the volitionist perspective, even more clearly a case of blamelessness than cases where the control/knowledge condition is missing but the attributionist can point to bad will. Others are more inclined to admit that the case is puzzling.

23. In the passage quoted above, it looks as though Calhoun is assuming a volitionist approach to understanding bad will, but we have no reason to assume that she would rule out an attributionist approach, and the example works well as a case where there is neither the volitionist type of bad will nor the attributionist. Calhoun's main concern is with the nature of the circumstances, what she calls "an abnormal moral context," (1989, 390) in which the behavior is wrong but widely thought to be permissible.

24. There is a different sense in which we hold people "responsible" without really believing that they are responsible: we assign liability. For example, we hold people responsible for what their dogs do. I am not talking about liability, but about responsibility, as will become clear.

25. Vargas (2013) argues that we should allow forward-looking considerations a more substantial role in our responsibility practices than we currently do.

26. See also Andrea Westlund's chapter in this volume for a discussion of the difference between holding responsible and blaming. The idea that we should take responsibility for things that are not under our control is not new. Joseph Raz, in his discussion of negligence (inadvertent fault), argues that we should take responsibility for negligence because our self-respect demands it (2010, 17). I discuss Raz's view in detail, as well as the literature on negligence, in Mason (2018, forthcoming). Robert Adams suggests in passing that we should take responsibility for traits that are not under our control (1985, 15). Susan Wolf suggests that there is a virtue of taking responsibility (2001, 10); David Enoch (2012) argues that we sometimes have a moral duty to take responsibility: I discuss his view below. There may be others that I have missed. Fischer and Ravizza (1998, 207–239) talk about taking responsibility, but it is clear that they intend for it to be apt only when the agent's act is under her control. Bernard Williams thinks that we are responsible for things that go beyond the boundaries of our wills, though he does not talk about *taking* responsibility (1993, 55–56).

27. See, e.g., Shoemaker's account (2015) of the various ways in which we might be impaired, and yet nonetheless have bad will.

28. It seems obvious that we couldn't be responsible in my sense for what others freely do (though we might be liable), but there is another interesting notion in this neighborhood, that I do not go into here, and that is "guilt by association" or "moral taint"—see, e.g., Oshana 2006.

29. See also Shoemaker 2015.
30. I defend this view in more detail in Mason 2015.
31. This is not supposed to be a complete account of respect. Obviously I am using an account such that respect involves affect. There is another sense of respect such that it is just a matter of treating someone appropriately, and in that sense I do think we have a moral duty.
32. For some different but not unrelated thoughts about feminist approaches to responsibility, see Oshana 2016.
33. A recent study (Extebarria et al. 2009) showing that women feel more guilt than men was much discussed on the Internet, and the discussion overwhelmingly has the tone, "tell me something I don't know."
34. First put into words by the sociologist Arlie Hochschild (Hochschild and Machung, 1989). For another relevant sociological study on gender differences in taking responsibility, see Walzer 1996.
35. Thanks to David Enoch, Alex Guerrero, Jen Morton, Monique Wonderly, and the editors of this volume for comments on earlier drafts.

References

Adams, R. M. 1985. "Involuntary Sins." *Philosophical Review* 94 (1): 3–31.

Arpaly, N. 2003. *Unprincipled Virtue: An Inquiry into Moral Agency.* Oxford: Oxford University Press.

Bargh, J. A. 2005. "Bypassing the Will: Toward Demystifying the Nonconscious Control of Social Behavior." In *The New Unconscious*, edited by R. R. Hassin, J. S. Uleman, and J. A. Bargh, 37–60. Oxford: Oxford University Press.

Bertrand, M., and S. Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market." NBER Working Papers from National Bureau of Economic Research, Inc., No. 9873.

Bessenoff, G. R., and Sherman, J. W. 2000. "Automatic and Controlled Components of Prejudice toward Fat People: Evaluation Versus Stereotype Activation." *Social Cognition* 18 (4): 329–353.

Brownstein, M. 2016. "Implicit Bias." *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/implicit-bias/.

Calhoun, C. 1989. "Responsibility and Reproach." *Ethics* 99 (2): 389–406.

Clarke, R. 2014. *Omissions: Agency, Metaphysics, and Responsibility*. Oxford: Oxford University Press.

Dasgupta, N. 2004. "Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations." *Social Justice Research* 17 (2): 143–169.

Enoch, D. 2012. "Being Responsible, Taking Responsibility, and Penumbral Agency." In *Luck, Value and Commitment: Themes from the Ethics of Bernard Williams*, edited by U. Heuer and G. Lang, 95–132. Oxford: Oxford University Press.

Etxebarria, I., M. J. Ortiz, S. Conejero, and A. Pascual. 2009. "Intensity of Habitual Guilt in Men and Women: Differences in Interpersonal Sensitivity and the Tendency Towards Anxious-Aggressive Guilt." *Spanish Journal of Psychology* 12 (2): 540–554.

Fazio, R. H., J. R. Jackson, B. C. Dunton, and C. J. Williams. 1995. "Variability in Automatic Activation as an Unobtrusive Measure of Racial Attitudes: A bona fide Pipeline?" *Journal of Personality and Social Psychol*ogy 69 (6): 1013–1027.

Fischer, J. M., and M. Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.

Frankfurt, H. G. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1): 5–20.

Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford: Oxford University Press.

Fricker, M. 2016. "Fault and No-Fault Responsibility for Implicit Prejudice—A Space for Epistemic Agent-Regret." In *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*, edited by M. Brady and M. Fricker, 33–50. Oxford: Oxford University Press.

Hochschild, A. R., and A. Machung. 1989. *The Second Shift*. New York: Penguin Books.

Holroyd, J. 2012. "Responsibility for Implicit Bias." *Journal of Social Philosophy* 43 (3): 274–306.

Kelly, D., and E. Roedder. 2008. "Racial Cognition and the Ethics of Implicit Bias." *Philosophy Compass* 3 (3): 522–540.

King, M. 2009. "The Problem with Negligence." *Social Theory and Practice* 35 (4): 577–595.

Levy, N. 2011. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press.

Mason, E. 2015. "Moral Ignorance and Blameworthiness." *Philosophical Studies* 172 (11): 3037–3057.

Mason, E. (2018, forthcoming) "Between Strict Liability and Blameworthy Quality of Will: Taking Responsibility." *Oxford Studies in Agency and Responsibility* 5.

McKenna, M. 2012. *Conversation and Responsibility*. Oxford: Oxford University Press.

Mills, C. W. 1998. *Blackness Visible: Essays on Philosophy and Race*. Ithaca, NY: Cornell University Press.

Oshana, M. 2006. "Moral Taint." *Metaphilosophy* 37 (3–4): 353–375.

Oshana, M. 2016. "A Feminist Approach to Moral Responsibility." In *The Routledge Companion to Free Will*, edited by M. Griffiths, N. Levy, and K. Timpe, 623–632. New York: Routledge.

Raz, J. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.

Rosen, G. 2002. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103 (1): 61–84.

Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.

Sher, G. 2009. *Who Knew?* New York: Oxford University Press.

Shoemaker, D. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.

Smith, A. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115 (2): 236–271.

Smith, H. 2011. "Non-Tracing Cases of Culpable Ignorance." *Criminal Law and Philosophy* 5 (2): 115–146.

Steinpreis, R., K. Anders, and D. Ritzke. 1999. "The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study." *Sex Roles* 41 (7–8): 509–528.

Strawson, P. F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48: 1–25.

Talbert, M. 2008. "Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy?" *Pacific Philosophical Quarterly* 89 (4): 516–535.

Tannenbaum, J. 2015. "Mere Moral Failure." *Canadian Journal of Philosophy* 45 (1): 58–84.

Vargas, M. 2005. "The Trouble with Tracing." *Midwest Studies in Philosophy* 29 (1): 269–291.

Vargas, M. 2013. *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press.

Wallace, R. J. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Walzer, S. 1996. "Thinking about the Baby: Gender and the Division of Infant Care." Social Problems 43 (2): 219–234.

Watson, G. 1987. "Responsibility and the Limits of Evil." In *Responsibility, Character, and the Emotions*, edited by F. Schoeman. New York: Cambridge University Press.

Watson, G. 1996. "Two Faces of Responsibility." *Philosophical Topics* 24 (2): 227–248.

Watson, G. 2004. "Responsibility and the Limits of Evil." Reprinted in *Agency and Answerability: Selected Essays*, 219–259. Oxford: Oxford University Press.

Watson, G. 2004. "Two Faces of Responsibility." Reprinted in *Agency and Answerability: Selected Essays*, 260–288. Oxford: Oxford University Press.

Williams, B. 1981. *Moral Luck: Philosophical Papers, 1973–1980*. Cambridge: Cambridge University Press.

Williams, B. 1993. *Shame and Necessity*. Berkeley: University of California Press.

Wolf, S. 2001. "The Moral of Moral Luck." *Philosophic Exchange* 31 (1): 5–19. http://digitalcommons.brockport.edu/phil_ex/vol31/iss1/1.

Zimmerman, M. J. 1997. "Moral Responsibility and Ignorance." *Ethics* 107 (3): 410–426.

AQ: Please provide page numbers for "Watson, G. 1987." (p. 35)