WHEN YOU MUST TAKE RESPONSIBILITY THOUGH YOU’RE NOT TO BLAME

Larisa Svirsky

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2014

Approved by:
Geoffrey Sayre-McCord
Susan Wolf
Matthew Kotzen
There are many situations where an agent’s responsibility for morally problematic behavior, attitudes, or aspects of her character seems to entail that she is blameworthy. There are some cases, however, where intuitively claims about responsibility and blameworthiness can and should come apart. In this paper, I argue that there is a sense of responsibility that agents have in virtue of having an obligation to take responsibility for behavior and features of their psychology that are attributable to them. In contrast, for agents to be blameworthy for some bad state they find themselves in, there must have been some reasonable expectation that they could have avoided that state. I illustrate this distinction between moral responsibility for something bad on the one hand and blameworthiness on the other with several cases, with the aim of arguing for additional subtlety in our evaluation of agents as morally responsible.
Introduction

There are many situations where an agent’s responsibility for morally problematic behavior, attitudes, or aspects of her character seems to entail that she is blameworthy. There are some cases, however, where intuitively claims about responsibility and blameworthiness can and should come apart. Below, I will describe several such cases, and argue that in spite of significant differences in these examples, the agents in question are responsible but not blameworthy for the same sort of reasons. In doing so, I aim to illuminate a sense of ‘moral responsibility’ that is often overlooked, thus allowing for additional subtlety in our evaluation of agents. I will discuss what the judgment that an agent is responsible in this sense indicates over and above the claim that it would be beneficial in some sense for her to take responsibility, and I will consider some cases where taking responsibility seems morally obligatory though one is not to blame. Cases where agents have a moral obligation to take responsibility for something that is attributable to them (i.e., something that expresses their evaluative commitments) strike me as cases where agents just are responsible in an important sense, so I use these phrases interchangeably.\(^1\)

---

\(^1\) It is not the obligation to take responsibility alone that entails that one is responsible. There are cases of strict liability where one is legally obligated to take responsibility for something without one actually having intended to harm anyone or been negligent, and there may well be analogous cases involving moral obligation. In at least some of these cases, it seems plausible that agents are not responsible, though they must take responsibility. The attributability of the state to the agent plays an important role in making it the case that she is in fact responsible for it because that relation establishes that the state itself says something important about her. I will say more about attributability as a necessary condition for responsibility in this sense later on in the paper.
Let us make a first pass at understanding moral responsibility without blameworthiness by considering the following case: Suppose the chancellor of a university has had an accounting error occur on her watch that costs the university a lot of money. She didn’t personally cause the accounting error, and she did everything one could reasonably expect her to do in hiring the appropriate staff and more generally overseeing the operations of the university. Although in principle the error could have been prevented, it is intuitively not appropriate to blame the chancellor for its occurrence, since she didn’t cause it, nor was she negligent in performing the tasks associated with her job. Still, in virtue of the kind of job she has, she doesn’t just incur duties to do what she can to remedy the accounting error – these are duties that many people might have regardless of their role in the situation; she also incurs a duty to take responsibility for the error when it happened. The appropriateness of her taking responsibility, moreover, does not derive solely from pragmatic reasons that follow from the circumstances she finds herself in. In other words, she is not just responsible in the sense that it would be good for the university if she took the heat for this unfortunate state of affairs. Rather, the role she plays in the university makes it the case that she ought to investigate with an aim of being able to explain how this accounting error could have occurred on her watch, and evaluate any university procedures that allowed it to happen. This is so regardless of whether there’s much she can do to remedy the situation going forward.

Because of the chancellor’s role within the university, she has a certain kind of standing with respect to the accounting error: she is answerable for it. By this I mean that her position involves certain standing duties regarding the goings on of the university, and those duties entail that when something goes wrong, she ought to be in a position to explain how the thing in question could have occurred, and evaluate whether that explanation justifies the current
university policies. It seems to me that this means that the chancellor has a kind of responsibility for the accounting error, though responsibility in this sense does not entail blameworthiness. For the chancellor to fail to take responsibility in such a case would be for her to ignore facts about her position and what it requires.

The Relevant Sense of Responsibility

I will argue that agents with the ability to explain and evaluate their own behavior, attitudes, and traits of character have a type of responsibility that is distinct from but analogous to the chancellor’s for those actions and features of their psychology; still, it might be the case that they are not to blame for them even when they are bad. Before saying more about specific examples where I will argue that responsibility and blame for bad things come apart, however, let me first clarify the senses of ‘responsibility’ and ‘blame’ I take to be relevant here. Of course, ‘responsibility’ has been used to mean a lot of things, in practice and in philosophical literature, and many of those senses of responsibility are not the one under discussion. For example, I am not discussing mere causal responsibility – the kind of responsibility something has in virtue of having caused something else. Of course, causal responsibility also comes apart from blame in many instances, particularly in cases where the causally responsible entity is not a moral agent at all, but this is not a particularly surprising claim.²

Additionally, when I say that an agent is responsible, I don’t just mean that we are entitled to hold her responsible for pragmatic reasons, where holding responsible involves, for example, asking for reparations or a resolution of the situation. There are many cases in which holding an agent responsible may be the best thing to do from an pragmatic perspective, and this

² Of course, ‘blame’ has also been used to mean a lot of things, and some kinds of blame might be appropriate to employ towards e.g., children, animals, and natural disasters. But there still seems to me to be a distinctive and central kind of blame that is only applicable to moral agents.
may be independent of whether or not she is blameworthy, but it may also be independent of whether or not she was actually responsible in either a causal sense or the sense I am interested in. Although I emphasized that the chancellor should take responsibility, I don’t just mean that it would be good for her to assume certain duties going forward, regardless of whether those duties are properly hers. Rather, as agents we have certain standing obligations because we have the ability to explain and evaluate our own behavior, attitudes, and traits of character. The existence of these standing obligations, like (but not the same as) the chancellor’s role obligation, entails that when things that we are involved in go wrong, we have to exercise those abilities, just as the chancellor has to investigate in light of the accounting error in order to be carrying out her job properly. When I say that these agents must take responsibility, then, I mean something stronger than the claim that it would be nice for these agents to assess themselves and their role in a situation; rather, I think doing so is required.3

Finally, I mean something a bit stronger than the claim that we can attribute bad actions to agents without thereby blaming them. This ‘attributability’ sense of moral responsibility typically involves seeing agents’ behavior and values as expressive of their evaluative commitments without thereby assuming that they had control over which behaviors and values they adopted as their own. As Gary Watson writes when characterizing this view,

> “These evaluations [of agents as responsible in the attributability sense] are inescapably evaluations of the agent because the conduct in question expresses the agent’s own evaluative commitments, her adoption of some ends among others…if what I do flows from my values and ends, [my activities] express what I’m about, my identity as an agent. They can be evaluated in distinctive ways (not just as welcome or unwelcome) because they themselves are exercises of my evaluative capacities” (Watson 233).

3 In some cases more than others it seems plausible to me that one has a duty to take responsibility that is a kind of moral obligation. In other cases, it seems more likely that taking responsibility would be beneficial, but it is supererogatory. I will discuss this distinction more later in the paper.
One thought behind attributability views is that no matter what causal path we follow to come to act or feel in the ways we do, those actions and feelings are still ours. Someone who has a terrible upbringing may well have not had much of an opportunity to develop certain good traits of character, and so it may well be inappropriate to blame him for the unfortunate character he in fact has. Nevertheless, Watson here points to a kind of evaluation that still seems appropriate in such cases because this agent’s character and actions are still his own, and expressive of his evaluative commitments.

This ‘attributability’ view of moral responsibility will be the most difficult to differentiate from my view, as I am discussing responsibility for actions and features of one’s psychology that are attributable to agents. I see attributability as a necessary condition for responsibility in my sense, but not a sufficient one. To see why this is the case, consider the example of someone with an eating disorder whose illness is at a stage where it is interfering with her capacity to explain and evaluate her own behavior. I think it is plausible to characterize her actions as still attributable to her in Watson’s sense. While this is a somewhat controversial point, I take it that her engagement in eating disordered behavior is expressive of something about her values. Nevertheless, if she is not in a position to answer for her behavior because of the state she is in, then I think she is not responsible for that behavior though it is attributable to her.

4 In the chancellor’s case, the accounting error is not attributable to her, for it is not expressive of her evaluative commitments. Thus, she is not responsible for the error because she has an obligation to take responsibility for it and it is attributable to her, but rather because she bears a relation that is analogous to attributability to the accounting error in virtue of her present role within the university. The purpose of the chancellor example is primarily to illustrate that in our existing practices, we sometimes hold people responsible for bad things though we think they are not to blame for them. There may be several ways to hold agents responsible in such circumstances, all closely related but not identical.
Whether or not one accepts this example as a case where an agent is responsible in an ‘attributability’ sense, my aim here is just to suggest that it would be helpful to develop a view like Watson’s in further detail, which I take myself to be doing here. What I aim to add to Watson’s view is a richer account of why having a morally significant property (e.g., being mean or being a person who caused harm) is something you have a duty to acknowledge in yourself, in the way I think we are obligated to take responsibility even in some cases where we are not to blame. It is not clear that attributability alone is enough to explain why being responsible for something is as ethically significant as it seems to be in the context of our practices. The additional element I think we should add here to an account of responsibility is that to be a responsible agent, you must have the ability to answer for those actions and features of your psychology that are attributable to you.

The value of answering for oneself in this way, and of taking responsibility in the sense I have tried to isolate, is the value of owning your behavior, attitudes, and character traits — in other words, seeing those actions and features of your psychology as saying something important about you. Whether you would have deliberately chosen to be an ungrateful person or a sexist, the importance of taking responsibility for those traits when you have them stems not only from the fact that they belong to you (as do all sorts of other traits that are not morally significant), but also that they are revelatory of your character. When the states you find yourself in are morally undesirable, perhaps even by your own lights, it is tempting to say that you ought not see them as expressive of your character. If you know, for example, that you have sexist attitudes and wish that you didn’t, it may be tempting see this desire to rid yourself of those attitudes as trumping

^What I say here about attributability owes a lot to Angela Smith’s discussion in “Attributability, Answerability, and Accountability: In Defense of a Unified Account” (2012), although our positive views about what answerability consists in are different.
the attitudes themselves in terms of your self-conception. Still, dissociating yourself from your behavior and psychology strikes me as inappropriate in such circumstances as well as unhealthy, and part of taking responsibility in the sense I have articulated involves preventing this kind of dissociation.

**What Being a Responsible Agent Entails**

Being a responsible agent (i.e., an agent who is potentially responsible for her behavior, attitudes, and character traits) in the sense I am interested in involves having two related abilities: the ability to explain your behavior, attitudes, or traits of character, and the ability to evaluate whether that explanation justifies those actions or features of your psychology. I take these general abilities to be fairly widespread. Being an agent of this kind then entails that you have certain standing obligations to exercise those explanatory and evaluative abilities, particularly when something you are involved in goes wrong, and to try to make amends for your wrongdoing. Having the general abilities that make a person a responsible agent need not entail that one is always able to exercise them successfully. Often when one acts in ways that are morally problematic, one is not immediately able to explain why one has done so, or even to see that one has done something wrong.

Still, responsible agents, in virtue of having certain standing obligations, must ask themselves why they’ve acted and felt in certain ways; failing to do so can lead to a kind of culpable self-ignorance about their moral standing. In the chancellor’s case, involving an analogous type of responsibility, the chancellor must investigate what features of the university’s policies made it possible for the accounting error to occur, and ask whether those policies are justified in light of the role they played in allowing an unfortunate state of affairs to take place. The chancellor may not be able to exhaustively answer those questions, but nevertheless, it is
part of her job to ask them in earnest. Similarly, I will argue, part of fulfilling our standing obligations as responsible agents involves examining some of the situations and states we are involved in and asking ourselves how we found ourselves there, and whether we ought to be there.\(^6\) When we find that we have done or felt something bad, moreover, we must do what we can to ameliorate the situation.

Blame, in contrast, has to do with the fact that often when we act and feel in ways that are morally problematic, we could have avoided doing so if we had tried to, and moreover we could have reasonably been expected to try. In other words, to be blameworthy, there must have been voluntary actions or omissions you could have reasonably been expected to perform such that you could have avoided the morally bad actions or features of your psychology that are attributable to you. When you (as a responsible agent) are responsible but not blameworthy for some particular thing, that means that you are in a position to fulfill your standing obligations to answer for yourself, but you could not have been reasonably expected to avoid the morally problematic thing that is attributable to you.

On my view, then, responsibility and blame are not wholly unrelated, but they are conceptually distinct. Being a responsible agent essentially involves the capacities we have to seek explanations for the way that we act and feel and to judge whether those explanations justify those actions and those feelings.\(^7\) In contrast, blameworthiness requires the ability to

---

\(^6\) Of course, examining all such states would send someone on a never-ending (not to mention narcissistic) quest to figure out exactly their place in the world. I take it that the states we find ourselves in that are most important to take responsibility for are the states that are most revelatory of our characters, as well as states that are of great moral significance.

\(^7\) By explanation, I don’t have anything particularly cognitively sophisticated in mind. The ability in question is roughly the ability to answer the question of why one has behaved or felt in the way one has. The only ingredient to these explanations that strikes me as essential is that they
avoid certain actions and feelings, and a reasonable expectation of that avoidance. As such, responsibility is a necessary condition for blameworthiness, for one could not have been reasonably expected to avoid behavior or features of one’s psychology that one was not in a position to assess, but it is not sufficient.

Being a responsible agent in the relevant sense also entails a standing duty to take stock of yourself periodically so that you are in a good position to successfully explain and evaluate morally significant features of your behavior and psychology. This taking stock of yourself involves investigating what your behavior, attitudes, and character are like, and if there is something you could reasonably be expected to do to improve those things. If you ignore this duty to try to know yourself, you can become culpably ignorant of things you could do to avoid bringing about bad states of affairs, such that if those states of affairs occur, you are to blame in virtue of the fact that you could have been reasonably expected to prevent them. Of course, not all instances of self-ignorance are culpable. If an agent has some reason why it is impossible or even simply quite difficult for her to acknowledge certain features of her psychology, she may be excused (at least to some degree) for the consequences of her self-ignorance, for it was not appropriate to expect that agent to know herself in that way. I will explore such a case later on in this paper.

One consequence of this view is that responsibility and blameworthiness are not all or nothing, for the abilities involved in answering for our behavior and features of our psychology and the extent to which we can be reasonably expected to have avoided bad actions, attitudes, are at a high enough level of description that they include the agent as such. Being able to give an explanation of some behavior at the level of particle physics would thus not suffice for having the ability to explain one’s behavior in the sense that I take to be relevant to moral responsibility. There is a sense, then, in which these explanations have to capture the sense that the behavior, attitude, or trait of character is attributable to an agent, whether or not she acquired them voluntarily.
and character traits clearly admit of degrees. This seems like a good feature in an account of responsibility, blame, and most other evaluative attitudes for that matter; after all, our evaluative judgments are often quite subtle, and sensitive to many features of situations agents find themselves in in a way that makes a gradational account of those attitudes seem appropriate. Moreover, the sense of responsibility I have tried to isolate requires agents to have certain abilities to explain and evaluate themselves, which are not plausibly abilities that we are born with, but rather ones that develop over time. This means that our judgments of responsibility should not only admit of degrees, but should also be judgments we make only with respect to those agents who have had the opportunity to develop the relevant abilities and not, for example, young children or people who are severely cognitively impaired. These predictions that my account yields about which agents are responsible and to what degree thus do seem to track some features of our practices of holding agents responsible.

Of course, we often talk about responsibility in closer relation to other evaluative attitudes, including praise and blame. I do not mean to deny that this is a legitimate use of the term, or one that is quite prevalent in ordinary language. Rather, I want to argue that there are some cases where blame seems inappropriate, but nevertheless we can meaningfully and appropriately ascribe responsibility to an agent. Additionally, it seems in some of these cases that agents have a duty to take responsibility, and if they do not fulfill this duty, they fail to recognize important facts about themselves. In some cases, this failure is itself blameworthy because it constitutes a kind of culpable self-ignorance.⁸

---

⁸ In other cases, it seems plausible to me that this failure to know oneself is unfortunate, but not itself blameworthy. At least some instances of psychopathology, which I go on to discuss later in the paper, seem to be cases where failing to know oneself in this way is unhelpful but understandable.
One might wonder what the judgment that an agent is responsible adds to our evaluation of her if we do not ascribe responsibility in order to justify praise, blame, or related attitudes. I believe that the ascription of responsibility in such contexts is a call for the agent in question to own her actions or features of her psychology, and to recognize their consequences for those around her. We can see the nature of this judgment most clearly if we look at some cases where agents are plausibly seen as responsible but not blameworthy. Below, I will consider three families of cases in turn: “involuntary sins,” certain instances of psychopathology, and finally certain instances of moral luck.

Revisiting “Involuntary Sins”

Consider a case of someone with a morally problematic attitude say, jealousy towards a friend of hers who recently received a promotion at work. She never directly expresses this attitude towards her friend, and moreover, she has done basically everything she can to avoid having such an attitude. She knows that she tends to be a jealous person, perhaps because of some facts about her upbringing, but she has gone to therapy to understand and ameliorate her jealousy towards others. Although our judgments about responsibility and blameworthiness in an example like this are not straightforward, the following claims about the case strike me as plausible: Even if she has exhausted what she can do about this attitude voluntarily, and even if it is not something she expresses to her friend, being jealous of her friend’s success is morally bad, and morally bad about her. Angela Smith expresses the significance of a more general version of this point about involuntary attitudes as follows:

“…it is a mistake to try to account for a person’s responsibility for her own attitudes in

\[\text{____________________________} \]

9 I follow Robert Adams in using this locution. Although he uses this term to refer to traits he describes as “inherently blameworthy,” I mean it only to refer to states (e.g., attitudes or traits of character) that are morally problematic but involuntary.
terms of their connection to her prior or future voluntary choices, because that obscures the special nature of our relation to our own attitudes: we are not merely producers of our attitudes, or even guardians over them; we are, first and foremost, inhabitants of them” (Smith 251).

Still, it seems inappropriate to blame the jealous friend for having this attitude; after all, she’s done more or less everything she can to rid herself of it. Still, her jealousy is attributable to her, and she should be concerned when she experiences it insofar as it is bad (both by her own lights and in general). In addition to the responsibilities she has to continue doing whatever she can to work on her jealousy, she should ask herself why she has this attitude and evaluate it, or in other words, ask herself whether or not it is justified.

According to my account of the relationship between responsibility and blame, then, it would be appropriate to describe the jealous friend in this case as responsible but not blameworthy for her jealousy. She is excused from blame in this case because her jealousy is not something she could have reasonably been expected to avoid; as I have described the case, she has exhausted all of the reasonable voluntary actions and omissions that would have allowed her to avoid that jealousy. It is not entirely uncontroversial, however, that an attitude’s being involuntary means that one cannot be blameworthy for it. In fact, Robert Adam and Angela Smith claim that voluntariness is not a necessary condition for either responsibility or blameworthiness. Adams argues that attitudes such as jealousy and self-righteousness are not morally bad only insofar as they come from our voluntary actions or omissions; rather, they are morally bad full stop. Moreover, Adams argues that we are blameworthy for those attitudes even if we have done whatever we can to avoid having them. In other words, blameworthiness extends beyond the bounds of the voluntary, and because we can be blameworthy for unfortunate attitudes or traits of character that we can’t seem to rid ourselves of, we must also be responsible for them.
In response to a view that he attributes to Lawrence Blum, according to which one can morally criticize someone for having racist attitudes, for example, but cannot properly blame him if they are not voluntary, Adams writes the following: “Perhaps for some people the word 'blame' has connotations that it does not have for me. To me it seems strange to say that I do not blame someone though I think poorly of him…Intuitively I should have said that thinking poorly of a person in this way is a form of unspoken blame. I am not sure, however, that there is only a verbal dispute here, about the use of the word 'blame' (Adams 21). Whether or not the dispute between Adams and Blum is merely verbal, I think there is an important distinction to be made between the claim that an agent is responsible for his “involuntary sins” in the sense I have articulated here and the claim that this agent is blameworthy. To get clearer about this distinction, we should consider another case.

Consider a person who is excessively ungrateful for his lot in life. Though he acknowledges that others have it worse, and he would like to feel more gratitude for the ways in which his life is going well, he finds himself instead resentful of little things that have gone wrong recently. He has done what anyone could reasonably expect him to do towards cultivating the gratitude he wishes he had, but his ingratitude remains recalcitrant. By Adams’ lights, this person ought to take responsibility for his ingratitude. For Adams, “The acceptance of responsibility is important to repentance because it enlists the desire to satisfy one's responsibilities in support of the desire to change; whereas if one says, ‘I'm not to blame for my ingratitude because I can't help it,’ one takes some of the pressure off oneself by seeking refuge in an excuse” (Adams 16).

According to Adams, moreover, blame is licensed in a case like this because it helps the individual to understand the wrongness of his involuntary states, which is important not only as a
recognition of the agent’s moral status, but can also be instrumental to the agents’ repentance and moral self-improvement. In addition to discussing the importance of the agent’s taking responsibility for such a state, Adams sees self-reproach (i.e., a form of blame) as an integral part of that practice. As he writes,

“…if you take responsibility for…[e.g., your ingratitude]…you also do not see it as something that just happens to you, like a toothache or a leak in your roof. You see it as an opposition that you yourself are making, not voluntarily but nonetheless really, to the generosity of the other person and to your own position as a recipient of love and assistance. In repentance you repudiate this opposition, not as an evil existing outside the inner circle of your selfhood, but as your own; and you reproach yourself for it.” (Adams 15-16).

I am unsure how deep my disagreement runs with Adams, but I would describe the case as follows: The ungrateful person, insofar as he is capable of asking himself for an explanation for his ingratitude, and asking whether that explanation justifies his having this attitude, is responsible for having it. Nevertheless, if he has really done all he can reasonably be expected to do to eliminate this trait in himself, he is not to blame for it. Recognizing that his ingratitude is attributable to him, engaging in the process of explaining and evaluating this trait, and doing what he can to be more grateful going forward are vital parts of his taking responsibility. It seems inappropriate for him to in addition feel guilt about this trait if he has worked sufficiently hard to avoid it, though it may well make sense for him to feel some negative emotion towards himself in virtue of his being responsible for a bad trait. Thus, if Adams’ insistence on self-reproach in a case like this involves claiming that the agent ought to feel guilt about the trait, there is a substantive difference between our views. If, on the other hand, Adams simply means that the agent must see the morally problematic trait as having something quite significant to do with him, and must do what he reasonably can do to avoid having it and exhibiting it, Adams and
I am basically in agreement about the appropriate response to cases like this, but have a verbal disagreement about the nature of ‘blame’.

It seems plausible that a similar kind of reasoning applies in the case of broader social attitudes, such as racist or sexist attitudes – that is, agents with racist or sexist attitudes who have done what they can to eliminate them (within reason) are responsible but not blameworthy for those attitudes. Their responsibility similarly stems from their general ability to explain and evaluate their attitudes, and they are excused from blame to the extent that they have worked to eliminate those attitudes in themselves. This explains an intuition I take to be commonly held, namely that agents who are raised in racist environments are typically taken to be less blameworthy for their racism than agents who are raised in more progressive environments. This seems plausible even when the agents acquired the attitudes via similar kinds of reasoning, there doesn’t seem to be a difference in whether the attitudes are attributable to the agents, and although the racist attitudes themselves are just as morally bad in either case. The explanation offered by my account is that your environment shapes what can be reasonably expected of you by making certain morally significant attitudes more or less salient and available. For someone

\[\text{\textsuperscript{10}}\text{This is not, of course, to say that someone who acts on their racist attitudes is not blameworthy for similar reasons. First of all, in the typical case of a person with explicit racist attitudes, they are blameworthy for those attitudes because we can reasonably expect agents to do more than they have done to eliminate them; after all, they are typically held in the face of massive evidence to the contrary. Secondly, it seems plausible to me that there are different standards for what we can reasonably expect others to avoid in the case of actions versus the case of attitudes, particularly attitudes that are held involuntarily.}\]

\[\text{\textsuperscript{11}}\text{For a related view, specifically concerning the connection between feminist critique, responsibility, and blame, see Calhoun (1989). An important part of her view is that in “abnormal moral contexts” where certain moral information (e.g., that certain practices are sexist) might be opaque to the general population, it may be very important to hold individuals responsible and in to some sense reproach them though they are not to blame due to the non-culpability of their ignorance. Holding agents responsible in such contexts though they are not}\]
raised in an overtly racist society, perhaps never encountering someone of another race, coming to hold anti-racist attitudes is a more significant moral and epistemic achievement than it would be for someone who has more exposure to the relevant information.\textsuperscript{12}

I see my view as a middle ground between the views of someone like Adams, who holds that we can be blameworthy for involuntary states and the view of someone who holds that voluntariness is a necessary condition for both responsibility and blameworthiness. It seems to me that we miss out on something important if we wholly excuse agents for morally problematic states that they ought to recognize in themselves, which they can understand and evaluate, and which they ought to make amends for exhibiting. Nevertheless, it seems excessively harsh to blame agents for actions, attitudes, and traits of character we couldn’t have reasonably expected them to avoid. If we avoid blaming the agents in question, we can still say something significant about them by saying that they are responsible for a something bad, namely that they ought to own certain evaluative commitments they have, whether or not they would have deliberately chosen them.

\textbf{Responsibility for Psychopathology}

Certain cases of psychopathology also strike me as cases where agents are responsible but not blameworthy for the behavior related to their illness. In general, mental illness is often blameworthy is an important part of making the relevant moral knowledge more accessible and thereby changing the moral landscape. As she writes, “…vulnerability to moral reproach is necessary to (1) publicizing moral standards, (2) conveying the obligatory force of moral commands, and (3) sustaining our sense of ourselves as self-legislators. In abnormal moral contexts, excusing excusable ignorance by withholding moral reproach inhibits the publicizing and adopting of new moral standards. Thus, in abnormal contexts, it may be reasonable to reproach moral failings even when individuals are not blameworthy” (Calhoun 405).

\textsuperscript{12} I set aside the question here of whether there are interesting differences between explicit racist and sexist attitudes and implicit bias and whether e.g., one can be responsible but not blameworthy for one’s implicit biases. My inclination is to say ‘yes,’ but to explore this question in the detail it requires would take me beyond the scope of this paper.
thought of as an excusing condition for responsibility as well as blame, but it seems to me that
the diversity of the phenomena in question calls for additional subtlety in our analysis of it.
Consider eating disorders, for example: people with eating disorders might well be in a position
to explain how they became ill and to know that the causes of their illness don’t justify their
engaging in eating disordered behavior. As long as their general explanatory and evaluative
abilities are intact, my account would have it that people with eating disorders are responsible for
their illnesses. Nevertheless, the difficulty in overcoming an eating disorder makes it
unreasonable to expect agents to consistently engage in healthy eating behavior early on in their
recovery.

Moreover, part of the symptomology of eating disorders includes a kind of ignorance
about the severity of one’s illness and one’s capacities to overcome it. The authors of a study
focused on the competence of anorexic patients to make decisions about their own treatment
write, “There appear to be two different types of belief described by participants: "factual belief,"
the belief of objective medical facts…and "salient belief," believing the information applies to
oneself in a way that is relevant to, and has an impact on, decision making” (Tan et al. 177).
These authors go on to say that these types of belief can come apart; in particular, although
eating disordered people (in this case, anorexic patients) can be quite good at judging abstract
facts about their situations, they have trouble integrating those facts into both their self-
conception and their decision-making. This lack of integration plausibly limits what people with
eating disorders can rationally intend to do, and what others can reasonably expect of them.
Thus, the difficulty involved in eating disorder recovery as well as this kind of self-ignorance
(that I take to be non-culpable) both function as excusing conditions for blame for a person’s
eating disorder.
Nevertheless, these considerations may not prevent people with eating disorders from being responsible for their illness in the sense I have tried to illuminate.\(^\text{13}\) The fact that someone is not yet in a position to consistently see her situation clearly need not entail that she is always ignorant of what she is doing when she acts, why she is doing it, or whether she ought to do it. People with eating disorders may well be able to explain and evaluate their behavior, and even see that this behavior causes suffering to themselves and others. Even if they do not engage in this kind of explanation and evaluation entirely successfully, moreover, many people with eating disorders can still ask themselves the relevant questions about the causes and justifications of their behavior. Finally, the practice of taking responsibility for one’s illness not only often plays an important role in one’s overcoming it, but additionally, failing to do so involves missing out on something important about one’s moral status.\(^\text{14}\)

This is not, of course, to say that people with eating disorders are always responsible for their illnesses. For example, in cases where agents are cognitively impaired because of malnutrition, they are not in a position to answer for their behavior, and thus they would not be responsible for it in the sense I have described. This seems to me to be the right result, since responsibility plausibly requires having and being able to exercise certain cognitive capacities.

---

\(^\text{13}\) These considerations might mitigate an agent’s degree of responsibility to the extent that they prevent her from seeing her behavior in context clearly enough to explain and evaluate it. The claim I am interested in establishing, however, is not that most people with eating disorders are responsible for their illness. Rather, I want to argue that, though people with eating disorders aren’t to blame for their illness, they can still be responsible for it insofar as they have the abilities I take to be central to responsibility in this sense.

\(^\text{14}\) One might think that failing to take responsibility when one has an obligation to do so is itself blameworthy. It is unclear to me that this is so in every case; in particular, in cases of mental illness, acknowledging one’s role in one’s own suffering (not to mention the suffering of those close to one) may be so painful that failing to do so would be unfortunate, but perhaps understandable and not blameworthy.
Additionally, in cases where agents are culpably self-ignorant, they might well be blameworthy for their actions because they ought to have acquired the knowledge they needed to carry out better behavior or hold better attitudes. The same holds in other cases where it would have been reasonable to expect agents to avoid the morally problematic behavior, attitudes, and traits of character they exhibit. It does not seem to me, however, that we are entitled to make this judgment in typical cases of psychopathology. People do not usually bring about their own suffering voluntarily, and the causes of mental illness are not usually within an agent’s control. Since blameworthiness requires not only some degree of voluntary control, but also the reasonable expectation of its exercise, one is not generally to blame for one’s having a mental illness.\footnote{15}

\textbf{Taking Responsibility When You’re Unlucky}

Finally, I believe it is appropriate to hold that agents are responsible but not blameworthy in certain instances of moral luck. The fact that much of what we do and how we are is not within our control, and the fact that we take ourselves to hold agents responsible (and potentially blameworthy) only for things that were within their control seems to entail that no one’s actually responsible for very much. And yet we do hold people responsible in spite of the apparent existence of moral luck: The child raised in Nazi Germany is in some sense unlucky that his environment made it more likely that he would come to accept Nazi ideology, and he did not choose to be born in that environment; still, now that he has accepted that ideology, he is intuitively at fault for doing so. Similarly, when circumstances outside of our control drastically

\footnote{15}{Although I only discuss eating disorders in detail in this section of the paper, I believe this account would apply to other cases of psychopathology that have a behavioral or attitudinal component just when the agent has the capacities necessary to answer for those aspects of her behavior and attitudes that constitute her illness.}
affect what we’ve done, regardless of our intentions, we still seem to be accountable for our actions. To quote Thomas Nagel, “…there is a morally significant difference between reckless driving and manslaughter. But whether a reckless driver hits a pedestrian depends on the presence of the pedestrian at the point where he recklessly passes a red light” (Nagel 26).

The issue of moral luck becomes even more pressing as a problem for our practices of holding agents responsible if we consider all agents who are morally responsible for bad things thereby blameworthy for those things. Above, I have offered reasons to think that this entailment between responsibility and blameworthiness does not hold, and I believe that separating responsibility from blame in the way I have described dissolves at least some worries related to moral luck. Below, I will outline two types of moral luck, constitutive and circumstantial luck (to use Nagel’s terminology), and illustrate them with examples that make clear how my account responds to such worries. First, in cases of constitutive moral luck, agents become the kind of people they are in virtue of being subject to forces outside of their control (e.g., their environments or upbringing). Although we typically hold people responsible (at least to some degree) for the kind of people they are as manifest in the things they do, it is not obvious that this is fair given that the way we acquire our characters is, in large part, not up to us.

Consider, for example, an agent who was raised in a household where excessive guilt and apology was expected and reinforced. A child growing up in such a household would be likely to develop dispositions to feel guilty and apologize in circumstances where doing so is not warranted. This trait is not only harmful to that individual insofar as it is a source of negative emotion, but it also irritates those around her. It strikes her friends as self-important that this person thinks she is responsible for bad events in the world that have basically nothing to do with her such that she thinks she needs to apologize for them. If we compare this agent to someone
with a similar personality but without a similar upbringing, however, we are less inclined to fault her for her tendencies towards excessive guilt, insofar as we see them as explained by the kind of environment that she was raised in. We thus see her as unlucky with respect to this disposition of hers; though it would be better if she didn’t have this disposition, it is unclear how to treat her in light of the fact that she does.

Part of this agent’s trouble is, of course, a tendency to take responsibility in excess for things in the world that she was not appropriately related to. Nevertheless, I want to argue that she ought to take responsibility for her dispositions to feel guilty and apologize, provided she is able to ask herself why she has those dispositions and whether she ought to have them, particularly when exhibiting them isn’t licensed by the circumstances she finds herself in. Of course, this agent is not to blame for those dispositions; her parents expected her to feel bad about everything that happened in their household, and to apologize even when she wasn’t at fault.\textsuperscript{16} It is understandable that she would continue exhibiting these feelings and behaviors into adulthood, and it may not be reasonable to expect her to do otherwise, depending on the difficulty of overcoming the influences of her upbringing. Having an agent take responsibility in these circumstances would involve her attempting to see why she has this tendency to feel excessive guilt, and to see that the reasons why she has this tendency don’t justify her continuing to have it. Even if there is little she can do to alter these tendencies in herself, having an awareness of the origin of her guilt and the fact that it is unjustified is part of recognizing an aspect of herself as a moral agent.

Compare this case with the following case of Susan Wolf’s: Suppose when visiting a friend’s home, someone accidentally breaks a vase. Let’s assume that she wasn’t particularly

\textsuperscript{16} To put this point more carefully, the agent is to blame for exhibiting these dispositions only insofar as she could have reasonably been expected to have avoided or overcome them.
careless in doing so (say, for example, the vase was sitting on the floor) and so it is relatively clear that she isn’t at fault for breaking the vase. There is a tendency, I think, to class the person who feels excessive guilt together with the vase-breaker as both instances where agents lack responsibility and are not to blame for their actions or feelings because they are involuntary. I want to argue that there is an important distinction to be made between the two cases: in the case of the person who feels excessive guilt as a result of constitutive moral luck with respect to her upbringing, she is obligated to take responsibility for her emotion because if she fails to do so, she misses out on something important about herself. In contrast, it would presumably be appropriate for the person who accidentally breaks a vase to take responsibility (and for example, apologize), but if she fails to do so, she is perhaps rude, but not self-ignorant. As such, her taking responsibility in such a case would be supererogatory.

My account also has consequences for cases of circumstantial moral luck where the agent’s circumstances affect the extent to which he can reasonably be expected to do otherwise. My account entails that in cases where agents could have been reasonably expected to do otherwise to the same degree, they are equally blameworthy, though differences in the consequences of their actions are reflected in differences in what they are responsible for. In classic instances of circumstantial moral luck where, for example, two drivers are equally negligent but one hits a pedestrian and one doesn’t, by my lights these agents are equally blameworthy because they are equally negligent. Moreover, they both ought to take responsibility for what they have done in the sense of attempting to explain and evaluate why they drove negligently. But in the case of the driver who hits someone, his taking responsibility also involves acknowledging what he alone has done, which was not just drive negligently but also hit a pedestrian. In virtue of his having hit a pedestrian, he incurs new responsibilities going
forward – perhaps, for example, to attempt to compensate the family of the person he hit. He cannot pretend that this event has nothing to do with him simply because he could have driven negligently and been lucky enough not to hit someone. Taking responsibility for something, then, involves not just factoring out the degree to which the consequences of our actions are up to us and accounting for those things we voluntarily chose to do. Rather, taking responsibility involves recognizing our place in the world as creatures with the ability to answer for actions and features of our psychology that are attributable to us, and that we find ourselves involved in whether we would have chosen to be so involved or not.

This view about taking responsibility in contexts where what we have done was in large part not up to us is consonant with Susan Wolf’s view in “The Moral of Moral Luck”. She writes,

“There is a virtue that I suspect we all dimly recognize and commend that may be expressed as the virtue of taking responsibility for one's actions and their consequences. It is, regrettably, a virtue with no name, and I am at a loss to suggest a name that would be helpful. It involves living with an expectation and a willingness to be held accountable for what one does, understanding the scope of "what one does", particularly when costs are involved, in an expansive rather than a narrow way…” (Wolf 121)

Although Wolf characterizes the disposition to take responsibility in this way as a virtue, I have here characterized it as a type of duty that one has at least in some cases where one is not to blame. I share her view, however, that insisting that we are only responsible for those things that are directly up to us is to draw a more or less arbitrary line around one’s voluntary actions and omissions and to mistakenly claim that one’s ethical life ends there. As she writes,

“The reason for objecting quite generally to an attitude of greater detachment and for commending an embrace of at least some of what lies beyond the sphere of one's will has less to do with a benevolent concern for others than with a view about what, for lack of a better word, might be called psychic health. The desirability of this [nameless virtue] comes partly from its expressing our recognition that we are beings who are thoroughly in-the-world, in interaction with others whose movements and thoughts we cannot fully control, and whom we affect and are affected by accidentally as well as intentionally,
involuntarily, unwittingly, inescapably, as well as voluntarily and deliberately…If we define ourselves in ways that aim to minimize the significance of contingency and luck, we do so at the cost of living less fully in the world, or at least at the cost of engaging less fully with the others who share that world” (Wolf 122-3).

Finally, in cases where an agent is entirely blameless for some unfortunate event she finds herself entangled in, and the occurrence of this event shows nothing about her evaluative commitments, she may nevertheless feel terrible for having been part of the causal chain that brought it about. Take, for instance, the case of someone who accidentally hits a child who has run in front of her car. There is significant disagreement about whether or not feeling terrible about one’s role in such a case (i.e., agent-regret) is appropriate or rational. Regardless, it seems clear enough that this is not an instance of responsibility without blameworthiness as I have described it in this paper. When the person who might potentially experience agent-regret attempts to answer for her behavior, she will find that though she was involved in an unfortunate state of affairs, this involvement shows nothing about her. Some evidence for this is the fact that in cases like this, it seems sensible enough to say things like, “It could have happened to anyone,” where this sort of thinking would undermine the appropriateness of taking responsibility in the sense I have outlined. Still, if a person did not feel any regret over accidentally hitting a child, that person would strike us as callous and perhaps even to blame for her lack of emotion. There may be an appropriate response to such circumstances that approximates taking responsibility for them, but this response is nevertheless distinct from fulfilling the obligations that follow from our evaluative commitments.

**Conclusion**

I have argued throughout this paper that there is an important sense of ‘moral responsibility’ that is appropriately ascribed to agents for bad actions, attitudes, and traits of character even when they are not to blame. In cases of “involuntary sins,” I have suggested
(contra Adams) that we are right to think that agents are not to blame for things outside of their control, though they still sometimes ought to take responsibility for them. In cases of psychopathology where agents have the general abilities necessary to answer for their behavior and features of their psychology, I have argued that this is enough for these agents to retain an important sense of responsibility that might well play an instrumental role in their recovery.

Lastly, when considering the problem of moral luck, my approach has been to agree with those who think that agents who had the same things within their control when they acted wrongly (e.g., two equally negligent drivers) are equally blameworthy, but nevertheless, they must take responsibility for the consequences of those actions even when those consequences are not those they would have chosen.

I think the sense of ‘responsibility’ I have discussed is one we recognize and use in ordinary practice, particularly in instances of role responsibility, as with the chancellor described above. Still, some of my suggestions about the relationship between responsibility and blame sound revisionary, even to my own ears, insofar as our practices of holding responsible are often bound up with other attitudes including praise and blame. I believe the cost of these revisions is well worth it, however, given the additional subtlety it buys us in terms of our evaluation of agents. I don’t mean to suggest that those who think responsibility and blame are often quite intimately connected are wrong; often agents with the ability to explain and evaluate their actions and features of their psychology could have reasonably been expected to avoid the bad states they find themselves in. Nevertheless, my aim here has been to show that even when agents meet excusing conditions for blame, they may well still be responsible, and ought to recognize that whether they would like to or not.


