

A Theory of Moral Blameworthiness

Pete Graham

September 29, 2007

In this paper I sketch an account of moral blame and blameworthiness. I begin by both clarifying what I take blame to be and explaining how blameworthiness is to be analyzed in terms of blame. I then consider different accounts of the conditions of blameworthiness. In the end, I settle on an analysis of blameworthiness according to which one is blameworthy for ϕ -ing just in case, in ϕ -ing, one violates one of a particular class of moral requirements that govern the attitudes we bear, and our mental orientation, toward other people and objects of significant moral worth. These requirements embody the moral stricture that we accord to these others a sufficient level of respect, one that their moral worth demands. This is a familiar theme which has its roots in P. F. Strawson's pioneering views on moral responsibility. My development of this Strawsonian theme leads me to the conclusion that the violation of a moral requirement governing action is *not* a condition of blameworthiness. On my account, violating a moral requirement to perform or refrain from performing an action is neither necessary nor sufficient for being blameworthy for something. All we are ever, strictly speaking, blameworthy for, I will show, are certain aspects of our mental bearing toward others. We can be said to be blameworthy for actions on my theory, only in the sense that those actions are the natural manifestations of the things for which we are strictly speaking blameworthy.

1. Blame and Blameworthiness

I take as my starting point the following platitude: X is blameworthy for ϕ -ing just in case X is worthy of blame for ϕ -ing. I take it, also, that X is worthy of blame for ϕ -ing just in case it would be appropriate for X to be blamed for ϕ -ing. All this is fairly uncontroversial. But what, exactly, is it for X to be blamed for ϕ -ing? As I see it, blame is, at its core, an emotional matter. That is, to say that X blames Y for ϕ -ing is to say, in the first instance, something about the emotions X feels toward Y in response to Y's ϕ -ing. It is true that one can say things like "I blame Jones for stealing the cookies" without feeling any particular emotions toward anyone. But performing such a speech act, it seems to me, serves the purpose, not so much to indicate that one stands in the blaming relation to Jones, but, rather, to indicate that one takes Jones to be blameworthy for

taking the cookies. But believing or taking Jones to be blameworthy for stealing the cookies is not the same thing as actually blaming him for doing so. What's more, believing someone to be blameworthy is not necessary for actually blaming her. One can blame someone for something while sincerely believing that she is not blameworthy for it—witness: “Inside, Mark just can't stop blaming Tanya for not calling him even though he now realizes that she was not at fault for not doing so.” Here, Mark's blaming Tanya is a matter of his having a certain emotional reaction to her, even though he does not think she is blameworthy for anything. It is evident from cases like this that standing in the blame relation to someone else just is a matter of having a certain emotional response to her, a response that may or may not fit with one's considered judgments about her blameworthiness.

As I will be understanding it, then, “X blames Y for ϕ -ing” is true just in case X feels one of a certain class of negative emotions toward Y in response to Y's ϕ -ing. As these emotions are constitutive of blame, I will refer to them as “the blame emotions”. But what emotions are these? Here, I will follow Jay Wallace (who, in turn, follows P. F. Strawson) in taking the paradigmatic blame emotions to be resentment, indignation, and guilt.¹ The first two are other-directed emotions, i.e., they are emotions that one feels toward others. Guilt, in contrast, is an emotion that is necessarily felt toward oneself.² Each of these emotions has the particular rebuking character constitutive of blame. Each might be considered an instance of anger, more generally, but their specific accusatory character sets them off as a class. This character also sets them apart from other classes of negative emotions. In particular, they are distinct from that class of negative emotions, characterized by shame and contempt, which have a comparative, rather than accusatory, character.

To resent someone for something, then, is to blame her for it. If Mary resents Bill for taking her ring, then she blames him for taking it. What is peculiar to resentment, among the blame emotions, is that the resenter views himself as a victim of the one he resents. More precisely, Y's resentment of X is Y's response to a perceived wrong or slight to Y on the part of X. Thus, Mary's resentment of Bill is her response to what she perceives to be Bill's wrongdoing

¹ In his seminal paper, “Freedom and Resentment” (Strawson 1962), Strawson first steered the discussion of moral responsibility toward an examination of a class of interpersonal attitudes and emotions he dubbed “the reactive attitudes”. In his book *Responsibility and the Moral Sentiments* (Wallace 1994), Jay Wallace refined this Strawsonian approach by insightfully focusing the discussion onto the three emotions I have here called the paradigmatic blame emotions and suggesting that blameworthiness be understood in terms of them.

² In the case of feelings of ‘collective guilt’, guilt is still reflexive, though it is felt toward one's group, and, in particular, toward oneself as a part of that group.

her when he takes her ring. Unlike resentment, indignation and guilt do not cast the one feeling them as the victim of the person toward whom they are felt. One can feel indignation toward someone even if there is no victim that that person has wronged or slighted. Similarly, one can feel guilty even if one does not see oneself as having wronged or slighted anyone. If Jones sincerely sets out to kill Santa Claus, it would certainly be appropriate for others to feel indignation towards him and for him, after reflecting upon what he has done, to feel guilty. So, whereas when Y resents X, Y's resentment is his response to X's wronging him, when Y feels indignation toward X, Y's indignation is his response simply to X's having committed a wrong, be it to someone in particular or not. Similarly, X's feeling guilty just is his response to his own commission of a wrong, again, be it to someone in particular or not. Thus, indignation and guilt are responses to wrongs, in general, whereas resentment is a particular response to wrongs to oneself.

These, then, are the blame emotions. They are the emotions the appropriateness the feeling of which toward someone is constitutive of that person's being blameworthy. In what follows, I will explore the conditions of blameworthiness in pursuit of a comprehensive theory. As this just is an exploration of the conditions under which the blame emotions are appropriate, I must first consider what, exactly, is meant by the claim that a certain blame emotion is appropriate.

2. The Correct Sense of "Appropriate"

From the preceding discussion we have the following analysis of blameworthiness:

X is blameworthy for ϕ -ing iff_{def} it is appropriate for a blame emotion—resentment, indignation, or guilt—to be felt toward X in response to X's ϕ -ing.

Among a number of questions that immediately arise about this analysis, one, in particular, jumps off the page: on what sense of "appropriate" is the appropriateness of feeling these emotions toward someone constitutive of her being blameworthy? The task of this section is to isolate and say more about this sense of "appropriate".

As a start in carrying out this task, it will be helpful to look, first, at one recent account. Wallace (1994) contends that the sense of "appropriate" relevant to the analysis of blameworthiness must be given a moral interpretation. Wallace takes his cue from the incompatibilist charge that it would be unfair to blame someone for ϕ -ing if her ϕ -ing was

causally necessitated by facts outside of her control.³ In this complaint Wallace sees the essence of the relevant notion of appropriateness: the feeling of a blame emotion toward someone for something is appropriate just in case it is not morally unfair that it be felt toward her for that thing. Thus, on Wallace's view, a person is blameworthy for ϕ -ing if and only if it would not be unfair for a blame emotion to be felt toward her in response to her ϕ -ing.

Can anything be said about this notion of fairness? Though Wallace does not offer much here, Gideon Rosen, an advocate of Wallace's account, says the following:

[t]he relevant notion of fairness is admittedly elusive. It is obviously not a matter of equal distribution of the reactive attitudes [i.e., guilt, resentment, indignation], or of procedural impartiality in their application. And yet despite this elusiveness it seems to me that we possess a robust body of opinion about when it is fair to treat someone adversely for what he has done. (Rosen 2002, 74)

Rosen is surely right that there is a notion of moral fairness that is both not a distributive matter and one about which we have a robust body of opinion. He is also right that the bulk of that body of opinion concerns the fairness of different forms of adverse treatment: it is a body of opinion about what types of punishment it is fair to impose on others for their transgressions. It is precisely because of this, however, that I am disinclined to think that this body of opinion has any relevance to the question of when the blame emotions are appropriate. As I have indicated, blaming someone is a matter of feeling certain emotions toward her. But if this is right, blaming someone is not a form of adverse *treatment* at all. Merely to feel certain emotions toward someone (even if they are negative emotions) is not to treat her in any way.⁴ (It might be true that X's blaming Y necessarily involves X's being *disposed* to treat Y in certain ways, but being disposed to treat someone in a certain way is not actually to treat her in that way, or in any way at all.) So, though there may be a body of opinion about when and where it is fair to punish people, it is not obvious how this body of opinion or this notion of fairness is relevant to when and where someone is blameworthy for something.⁵

³ Incompatibilists hold, whereas compatibilists deny, that, necessarily, if universal causal determinism is true, then no one is blameworthy for anything.

⁴ Even Wallace grants this: "Though blame often and naturally finds expression in sanctioning behavior, it is not necessarily so expressed—thus I can blame a person "privately" without expressing my response to anyone at all, much less sanctioning the person whom I blame (who may, anyway, be outside my sphere of causal influence)." (Wallace 1994, 56)

⁵ In a recent paper, Hieronymi (2004) has argued that blaming someone for something cannot be unfair. Like in my discussion, an important part of her argument is that blaming someone does not entail treating her in any particular way.

There also seems to be good reason to think that Wallace's fairness construal is false. If it is ever unfair to feel a blame emotion toward someone it must be because doing so is somehow bad for him.⁶ But there are some for whom, it seems, the mere feeling of these emotions toward them wouldn't be bad at all. Though most people don't want to have the blame emotions felt toward them, this needn't be true for all persons. It is not hard to imagine people who are completely indifferent to whether others feel such emotions toward them. (Imagination, here, might not even be necessary.) As feeling the blame emotions toward these people surely is not bad for them, it isn't clear that it would be morally unfair to feel these emotions toward them even when they aren't blameworthy for anything.⁷ Indeed, we can imagine people who greatly desire our indignation and resentment. People who repeatedly turn themselves in to the police for crimes they haven't committed may be real life examples of this. Now, it certainly would be very hard for us to muster up the indignation that they crave. But, in the event that we were able to get ourselves to be indignant toward them, it certainly wouldn't be *morally unfair* of us to do so. Nonetheless, in such a case, it would not follow from the fact that it was not morally unfair of us to blame them that they would be blameworthy for anything.⁸ The sense of "appropriate" relevant to the analysis of blameworthiness, then, should not be understood in terms of moral fairness.

I am inclined to think that the notion of appropriateness central to the analysis of blameworthiness is not to be understood in moral terms at all.⁹ My reason for this, however, is not just that there is no particular moral notion that I can think of that fits the bill. More important than this is that the character of debates over a person's blameworthiness do not seem to have the kind of moral flavor that they would have to have were it to be understood in such a

⁶ Note, here I am only contending that for the feeling of the blame emotions toward someone to be *unfair* it seems that the feeling of them toward him must be bad for him. I am not here claiming that for it to be *wrong* to feel these emotions toward someone, doing so would have to be bad for him.

⁷ It might be wrong of us to feel these emotions toward them if it was essential that in feeling these emotions toward them, we necessarily bore them ill will. But it is essential neither to resentment, nor to indignation, that feeling it toward someone entails bearing him any ill will.

⁸ I do not mean, here, to be affirming that a thing can be bad for someone only if that person does not want it to happen. I mean only to be affirming that when it comes to the blame emotions, if feeling them toward someone is bad for him, it is so only because they are negative emotions that he prefers others not to feel toward him. In the case of people who are indifferent to, or even desire, our feeling these emotions toward them, it doesn't seem bad for them if we do feel those emotions toward them. Thus, even if they are not, in fact, blameworthy for anything, it doesn't seem unfair of us to feel the blame emotions toward them. (It might be bad for them were we to treat them adversely, but, as I have indicated, blaming someone is not a matter of treating him in any particular way at all.)

⁹ In rejecting a moral interpretation of the notion of appropriateness central to the analysis of blameworthiness, I believe I am in agreement with D'Arms and Jacobson (2000) who argue that the appropriateness of an emotion is often (perhaps always) not a moral matter.

manner. To be sure, debates about whether X is blameworthy for ϕ -ing often do concern moral matters. But the focus of the moral dispute in these debates is the moral status of X's ϕ -ing, not the emotions we are inclined to feel toward him in response to it. So, though whether X's ϕ -ing is morally wrong certainly is germane to the question of whether she is blameworthy for ϕ -ing, facts, be they moral or not, about prospective blamers are not. In cases in which it is in dispute whether X is blameworthy for ϕ -ing, the defense's case is not that blaming X for ϕ -ing would be somehow morally objectionable, but that blaming X for ϕ -ing would be *a mistake*. X's blamelessness consists in the fact that those who blame X for ϕ -ing are in error, rather than themselves committing a moral offense. Even in cases in which it is in dispute whether X is blameworthy for ϕ -ing despite its being agreed that her ϕ -ing was morally wrong—as when X's ϕ -ing is done from blameless ignorance, or when X is a small child or is suffering from some psychic debilitation—what is at issue is whether blaming X for ϕ -ing would be correct, not whether it would be somehow morally objectionable.

I do not mean to be saying that it is never morally wrong to blame someone when she is not blameworthy. It may very well be that in some cases it is morally wrong to feel the blame emotions toward someone who is not, in fact, morally blameworthy for anything. All I am claiming is that a person's not being morally blameworthy *does not consist in* that moral wrongness. The person's blamelessness, I contend, consists in the blame's being mistaken or incorrect. So, it might be that there is a moral requirement that one not blame someone for ϕ -ing if in doing so one's blame would be mistaken. (I am skeptical that there is such a moral requirement, but I do not rule it out.) Even so, the question of whether a person is blameworthy for something would still not *itself* be moral in the way that Wallace and Rosen take it to be. As evidence of this, consider that people often offer as an explanation of *why* it would be morally objectionable to blame someone for something that she is not blameworthy for it. If its being morally objectionable to blame her *just was* her not being blameworthy, then this would be no explanation at all.

If the appropriateness of feeling a blame emotion toward someone is a matter of the blame's being, in some sense, correct, rather than morally unobjectionable, then that appropriateness must be a kind of factive appropriateness. Blame felt toward an object imputes to its object certain features, and it is to such features that blame is appropriately responsive. Any particular instance of blame, then, is appropriate just in case the object of that blame, i.e., that

toward which it is felt, has the features that the blame emotion imputes to it. Blame is like many other emotions in this regard. Take fear. Fear imputes to its object the property of being dangerous to its subject. Thus, when X fears a dragonfly, X's fear imputes to the dragonfly the property of being dangerous to X. This does not amount to a belief that the dragonfly is dangerous. The person who fears the dragonfly despite her recognition that it is innocuous is not irrational in virtue of believing of the dragonfly both that it is and is not dangerous. It's just that, though one of her mental states, one of her beliefs, has the content that the dragonfly is not dangerous, another of her mental states, her fearful experience, has the content that it is.¹⁰

Though there exists no word "fearworthy" in English, it is clear that there is a sense of "appropriate" that, as applied to fear, tracks the correctness of particular instances of fear. A particular fear is appropriate, in this sense, if and only if the propositional content of that fear, namely, that the thing feared is dangerous to the subject, is true. The appropriateness of a particular fear is clearly not a moral matter. Whether something is fearworthy depends not on moral norms governing fear, but on the correctness of the propositional content of the emotion. Now, this is not to deny that there are moral norms governing fear. Maybe there are such norms. It's just that the question of whether one's fear is appropriate is not a moral matter; it is just a matter of whether the propositional content of that fear is true. I take the situation as regards blame and blameworthiness to be analogous.¹¹

If the relation between blame and blameworthiness is analogous to that between fear and fearworthiness, the blame emotions have a particular propositional content. What's more, if this is right, then a person is blameworthy for ϕ -ing just in case the content of some blame emotion felt toward her in response to her ϕ -ing is (or would be) true. This is fine so far as it goes. But what, exactly, is the content of a blame emotion felt toward someone in response to his ϕ -ing? Answering this question will complete my sketch of a theory of blameworthiness.

3. The Content of the Blame Emotions

¹⁰ The view that emotions have propositional contents goes by the name cognitivism in the philosophy of emotions. For a sketch of cognitivism see Roberts (1988) and Deigh (1994).

¹¹ Though I take the appropriateness of a blame emotion relevant to the analysis of blameworthiness to be a factive appropriateness, I take it to be factive only in the sense that the emotion is appropriate just in case the propositional content of the emotion is true. As will be clear in what follows, I do not hold that the contents of the blame emotions are factual in the sense that they are non-moral or non-normative. On the contrary, as I will argue, the blame emotions have a very particular kind of moral content.

So far, I have argued that a person's blameworthiness for ϕ -ing is to be understood in terms of its being appropriate that certain blame emotions be felt toward her in response to her ϕ -ing. I have further argued that this appropriateness is a matter of the propositional content of those emotions being true. Call this "the true-blame-emotion conception of blameworthiness". At this point, the following question naturally arises: what is the content of a blame emotion? The best way to answer this question, I think, is to consider candidate accounts and test them against our intuitive judgments about particular cases.

I begin with the following simple account:

Wrong Behavior: the content of a blame emotion felt toward X in response to X's ϕ -ing is that, in ϕ -ing, X has behaved wrongly.

Though **Wrong Behavior** is at first blush quite intuitive, it fails for at least three different reasons. First, a person can act wrongly and yet not be blameworthy. Second, a person can be blameworthy even though she does not act wrongly. Third, and most important for my purposes, it is not the case that a person is blameworthy for ϕ -ing only if her ϕ -ing is an instance of her behavior; that is, a person can be blameworthy for something other than her performing or failing to perform certain actions.¹² I will establish these points in order.

That it is possible for someone to act wrongly but not be blameworthy is the easiest to establish. It is a commonplace that there are instances of blameless wrongdoing. One easy way in which blameless wrongdoing can come about is through an agent's ignorance. Take the following case:

SUGAR: Unbeknownst to Bill, Sally has replaced the sugar in his sugar bowl with arsenic. As a result, Bill unwittingly poisons his neighbor when she comes over for a cup of tea.

In SUGAR, Bill acts wrongly. He poisons his neighbor by spooning arsenic into her tea. In doing so he lethally poisons his neighbor, and he morally ought not to lethally poison an innocent person. Though in spooning arsenic into his neighbor's tea he behaves wrongly, he is not blameworthy in doing so. SUGAR, thus, is a paradigm case of blameless wrongdoing.

¹² I am using "behavior" here, perhaps non-standardly, in such a way that a person's behavior consists of both those actions he performs and those actions he fails to perform. The important point, for which I will argue, is that a person can be blameworthy for things other than his performing or failing to perform certain actions, and my preferred succinct way of saying this is to say that a person can be blameworthy for non-behavior.

That it is possible to act morally permissibly and yet nonetheless be blameworthy is a little less obvious. Like cases of blameless wrongdoing, such cases can arise in virtue of an agent's ignorance. Take the following case:

UNWITTING DEFENDER: Unbeknownst to Bob, Sue is about to detonate a bomb that will kill a thousand innocent people. The only way to prevent her from doing this, given the situation, is to kill her. Bob very much dislikes Sue and, for his own personal reasons, pulls out a revolver and shoots her dead.

In this case, in shooting Sue, Bob acts morally permissibly. It is certainly morally permissible to kill one person if failing to do so would result in her villainously killing a thousand innocent people. Though he acts permissibly in this case, Bob is, it seems, nonetheless, blameworthy. Indignation toward him would certainly be appropriate. Had Bob known that Sue was about to kill a thousand people and shot her in order to prevent her from doing so, then indignation would not be appropriate, and, accordingly, he would not be blameworthy. But because he has no idea that his killing Sue would have that good consequence, he is an appropriate target for our indignation.

Ignorance works in opposite directions in SUGAR and in UNWITTING DEFENDER. Whereas in SUGAR, Bill's ignorance is exculpatory, in UNWITTING DEFENDER, Bob's ignorance is incriminatory. Because of this, it might be thought that the content of a blame emotion is that the one toward whom it is felt has done something *subjectively* wrong. An action is subjectively wrong, roughly speaking, just in case it would be wrong were the agent's beliefs about her situation correct. One might think that switching to subjective wrongness can solve both of the problems for **Wrong Behavior** that I have raised. In SUGAR, it is not subjectively wrong of Bill to spoon the arsenic into his neighbor's tea. This is because were there no arsenic in the sugar bowl, as he blamelessly believes to be the case, his action would not be morally wrong. And in UNWITTING DEFENDER, it is subjectively wrong of Bob to kill Sue because were things as he took them to be—namely, that Sue was not about to kill a thousand innocent people—his action would have been morally wrong. Thus, one might suggest the following as a replacement for **Wrong Behavior**:

Subjective Wrong Behavior: the content of a blame emotion felt toward X in response to X's ϕ -ing is that, in ϕ -ing, X has behaved subjectively wrongly.

This, it might be thought, handles both SUGAR and UNWITTING DEFENDER quite nicely.

The shift to **Subjective Wrong Behavior** will not work, however, because there are counterexamples to it that do not involve any ignorance of the facts:

TROLLEY: James stands at a switch that, if he throws it, will divert a runaway trolley onto a side track, thereby preventing it from killing five trapped track workers. If the trolley is diverted onto the side track, however, it will run over and kill Chris, another trapped track worker. James is a philosophy professor who specializes in ethics. He understands all the morally relevant features of the situation and correctly concludes that it is morally permissible to divert the trolley onto the side track. He diverts the trolley, however, not to save the five people from being killed by the trolley, but to kill Chris. Chris is his despised uncle from whose death he will benefit. He flips the switch and, thus, kills Chris in order to inherit his fortune.¹³

Like Bob in UNWITTING DEFENDER, James acts in a morally permissible way. I take it that the trolley scenario is one in which it is morally permissible (though not obligatory) to kill one, if in doing so five others are saved from being killed. What's more, James understands this and so it is not the case that he acts subjectively wrongly. But even though he acts morally permissibly in turning the trolley and thereby killing Chris, James is, nonetheless, blameworthy.¹⁴ Chris's deep resentment of him in the last moments of his life surely would be appropriate, on any sense of "appropriate". Similarly, the indignation that others feel toward James would also certainly be appropriate. And, in the event that after much reflection and soul-searching James undergoes a reformation of character, the guilt he would feel would, again, certainly be appropriate.

So far, I have argued that **Wrong Behavior** is false by offering a number of scenarios as counterexamples to it. I have also shown that appealing to the notion of subjective wrongness, as embodied in **Subjective Wrong Behavior**, will not help. But **Wrong Behavior** and **Subjective Wrong Behavior** fail for another, more enlightening, reason: they both entail that a person can be blameworthy only for her behavior; that is, they both entail that a person can be blameworthy only either for her actions or her failures to act. In the next section, I show that this is false by illustrating, with a few examples, the breadth of the non-behavior for which a person can be blameworthy.

¹³ Here I borrow a version of the trolley scenario first introduced in (Kamm 2001). Kamm's purpose in proposing this case is to illustrate that irrespective of James's malign intention, it is just as morally permissible for him to turn the trolley in this scenario as it would be for someone else with a more benevolent intention to turn it. As is clear from the text, on this point, I entirely agree.

¹⁴ Though I contend that he is blameworthy for something in this scenario, as will be evident later in the chapter, *what* it is that he is blameworthy for, I will maintain, is not any action of his. None of this affects the point I am making here.

4. Blameworthiness for Non-Behavior

People can be blameworthy for things other than performing or failing to perform certain actions. One can be blameworthy for certain episodes in one's psychic history. Consider the following case:

TORTURE: Sam is being tortured mercilessly by Nikita. Adam, on his way home from work, passes a window into the room in which Nikita is torturing Sam. As Sam writhes in pain from the electric shocks applied by Nikita, he spies Adam peering in at the procedure. Sam knows that Adam knows that he cannot intervene and help Sam—Nikita has locked the door and the window's glass is unbreakable—and he also knows that Adam believes that Sam cannot see him—Sam knows that Adam is under the mistaken impression that the window through which he is looking is a one-way mirror. Amidst the excruciating pain, Sam notices that as the voltage is applied, and, thereby, his agony increases, there creeps across Adam's face a smile of sheer delight. At this sight, resentment toward Adam wells up within Sam.

In this case, it seems that Sam's resentment of Adam is entirely appropriate; there is no sense of "inappropriate", for that matter, on which it seems inappropriate. However, it is not the case that that for which Sam resents Adam is an action. One might say that in this case Sam resents Adam for smiling. But smiling is not an action that Adam performs. Rather, like blushing, it is more of a bodily occurrence than it is an action.¹⁵ Neither is Adam's smiling a failure of his to perform an action. In truth, however, it is *not* for his smiling that Sam resents Adam. What Sam resents Adam for is his taking pleasure in Sam's pain. Adam's smile is only the outward manifestation of that for which Sam really resents him. Taking pleasure in someone's pain, however, is clearly not something one does, in the sense of performing an action. Therefore, TORTURE demonstrates that a person can be appropriately resented, and, thus, can be blameworthy, merely for enjoying someone else's pain.

Not only can one be blameworthy for having certain experiences, one can also be blameworthy for having certain hopes and desires. Here is one such case:

EVIL HOPE: Prior to the events in TORTURE, while being led to the torture chamber by Nikita, Sam sees Eve alone in a room hoping aloud to herself that Sam's ordeal will be extremely painful. He knows that she believes that she is all alone and that no one can hear her. At this sight, resentment toward Eve wells up within Sam.

¹⁵ It is true that sometimes we can perform the action of smiling, but there are cases of spontaneous smiling that are not actions. It is this kind of spontaneous smile that occurs in Adam in TORTURE.

In this case, it seems that Sam's resentment of Eve is perfectly appropriate. There is certainly no sense of "inappropriate" according to which it is inappropriate. But, like in the case of TORTURE, that for which Sam resents Eve is neither an action nor a failure to perform an action. What Sam appropriately resents Eve for is her hoping for and wanting his suffering. Thus, EVIL HOPE demonstrates that one can be blameworthy for what one wants.¹⁶

My final example of blameworthiness for non-behavior illustrates that one can be blameworthy for certain aspects of one's deliberations:

ICE CREAM: Brad stands in front of an ice cream counter deciding between ordering mint chip and vanilla. He likes both flavors very much and knows that he will enjoy whichever one he orders. The man behind the counter informs Brad that if he does not pick mint chip, a group of thugs will savagely beat Ray, a defenseless teenager strapped to a chair in the corner of the room. Brad takes this information in and continues to consider his choice. After much deliberation, he chooses mint chip and Ray is thus unharmed. However, it is clear to Ray that in Brad's deliberations (Brad always mumbles his deliberations aloud) the only consideration that inclined him toward mint chip over vanilla was the fact that he had vanilla the last time he ate ice cream. Upon realizing this, there wells up within Ray a deep resentment toward Brad.

In this case it seems that, for any sense of "appropriate", Ray's resentment of Brad is entirely appropriate. But, again, what he resents Brad for does not seem to be an action. What he resents him for is the fact that he did not take Ray's well-being to be a relevant and significant factor in his deliberations about what to do. If this is right, then what Brad is blameworthy for is his failing to take Ray's well-being into account in his deliberations. But his failure to do so is not an instance of his failing to perform an action. Were Brad to take Ray's well-being into account in his deliberations it would not be something he did, in the sense of performing an action. If he took Ray's well-being into account, this would just have been a matter of Ray's well-being seeming relevant to him and, thus, causing the course of his deliberations to proceed in a different way than it in fact did. Similarly, Brad's not taking Ray's well-being into account just is a matter of Ray's well-being not seeming important to him at all, and, thus, simply not playing a role in influencing the course of his deliberations. So, in ICE CREAM Brad is blameworthy for failing to take Ray's well-being into account in his deliberations, but his failing to do so is not a matter of his performing or failing to perform some action.

¹⁶ In a recent paper, Angela Smith (2005) has also argued that people can be morally responsible for some of the propositional attitudes that they bear toward others.

A proponent of the view that one can only be blameworthy for behavior might suggest that what Adam, Eve, and Brad are blameworthy for in these cases are not their particular experiences, desires, deliberations, etc., but, rather, for actions and omissions of theirs in the past that resulted in their being the way they presently are. Adam, for example is not blameworthy for taking pleasure in Sam's pain, but, instead, for performing (failing to perform) those actions in the past such that had he refrained from performing them (had he performed them) he would not have been so horrible a person as to take pleasure in another person's excruciating pain.

This response fails for a couple of reasons. First, it implies that the blame emotions would be inappropriate if there were, in fact, no such actions or omissions of theirs in the past. The reply has it that had Adam made no such character-transforming actions or omissions in the past, Sam's resentment of him would be inappropriate. This, however, is highly implausible. For one thing, Adam's enjoying Sam's pain might not be something that is the result of any of his actions. And, for another, though there may, perhaps, have been certain actions he could have performed in the past such that had he performed them he would not have been the type of person he presently is, we needn't suppose that at any time prior to seeing Sam being tortured he knew or had any reason to believe that he was the type of person who would take pleasure in another person's pain. But if this is the case, how can it be that he is blameworthy for having failed to perform one of these character-changing actions, given that at no point in the past did he have any inkling of the fact that his character was in need of changing? All of this notwithstanding, it would still be appropriate for Sam to resent Adam if, in looking in at the torture session, he truly enjoys Sam's pain.¹⁷

Second, this reply conflicts with the phenomenology of the blame emotions. Sam's resentment of Adam is a response to features of Adam's present self, namely, his particular experience of enjoying Sam's pain, and not any behavior of his in the past. Sam's resentment is

¹⁷ These considerations, it seems to me, also tell against a slightly different thesis, advocated by Robert Audi (1991), namely, the Traceability Thesis. According to this thesis one is blameworthy for ϕ -ing only if either ϕ -ing is a basic act or there is some basic act for which one is blameworthy and it is in virtue of one's being blameworthy for that act that one is blameworthy for ϕ -ing. (The Traceability Thesis differs slightly from the proposal in the text because unlike the proposal in the text, the Traceability Thesis allows that it is possible for someone to be blameworthy for non-behavior.) If, as I have claimed, Adam would still be blameworthy for enjoying Sam's pain even though at no time prior to his seeing Sam being tortured he had any reason to believe that he was the type of person who takes pleasure in someone else's pain or that he needed to take steps to prevent himself from doing so, then, it seems, the Traceability Thesis is false. Adam is a counterexample to it, for his taking pleasure in Sam's pain both is not a basic act and is not such that he is blameworthy for it only in virtue of his being blameworthy for some prior act or failure to act.

clearly a response to something recent rather than remote. Realizing that he enjoys Sam's suffering does not play for Sam a mere evidential role in his resenting Adam: it does not play the role of establishing conclusively that Adam has behaved wrongly in the past. Rather, Adam's current enjoyment of his pain is transparently the object of Sam's resentment.

It will not do, then, to hold that, in TORTURE, EVIL HOPE, and ICE CREAM, what Adam, Eve, and Brad are blameworthy for are instances of behavior. Rather, it must be accepted that they are blameworthy for things other than their behavior.

5. Toward A New Account of Blameworthiness

So far, I have established that people can be blameworthy for things other than their behavior. If this is right, then, given that blameworthiness just is a matter of the truth of the content of the blame emotions, the content of a blame emotion felt toward an agent for ϕ -ing cannot be that in ϕ -ing the agent *behaves* wrongly. But now a question arises. Must we give up on the intuitive thought that the blame emotions are responsive to wrongdoing? Insofar as we mean "doing wrong" by "wrongdoing", where "doing" implies behaving in some way, it seems we must. But, insofar as we understand wrongdoing in a more expansive way, we needn't.

If someone behaves wrongly, then she either performs a morally prohibited action or she fails to perform a morally required action. If all moral requirements are requirements concerning behavior, then all wrongdoing consists either in an agent's acting or failing to act in some way or other. But why should we think that moral requirements only ever require people either to perform or refrain from performing certain actions? It seems arbitrary and dogmatic to restrict the application of moral requirements simply to actions and omissions. If some moral requirements are not requirements concerning action, then an agent can violate a moral requirement otherwise than by performing or failing to perform some action. If we allow "wrongdoing" to cover all violations of moral requirements, we can hold on, as I think we should, to the intuitive thought that the blame emotions are responses to wrongdoing. All we need admit is that not all blameworthiness is blameworthiness for some type of behavior. Thus, instead of **Wrong Behavior**, perhaps we should accept the following:

Wrongdoing: the content of a blame emotion felt toward X in response to X's ϕ -ing is that, in ϕ -ing, X has violated a moral requirement.

This account in conjunction with the true-blame-emotion conception of blameworthiness nicely accommodates the fact that people can be, and often are, blameworthy for things other than their behavior.

Which moral requirements, according to **Wrongdoing**, are violated in TORTURE, EVIL HOPE, and ICE CREAM? In TORTURE, recall, I argued that that for which Sam appropriately resents Adam is his taking pleasure in his pain, and so in that case Adam is blameworthy for taking pleasure in Sam's pain. According to the theory presently on offer, this will be true only if Adam is under a moral requirement not to take pleasure in Sam's pain. In EVIL HOPE, I argued that Eve is blameworthy for wanting and hoping for Sam's suffering. Again, if **Wrongdoing** is correct, this will be true only if Eve is under a moral requirement not to want Sam to suffer. And if I am correct that in ICE CREAM Brad is blameworthy for not taking Ray's well-being into account in his deliberations, then **Wrongdoing** is correct only if Brad is morally required to take Ray's well-being into account in his deliberations. **Wrongdoing** will satisfactorily account for these cases only if there are moral requirements of these sorts. If not, then the theory is in trouble.

The theory is not in trouble, however, for it seems clear that people *are* under moral requirements of these sorts. In general, people are morally required not to want or enjoy the suffering of others and they are morally required to take the interests of others into account when considering what to do. The following are perfectly ordinary things to say: "you really ought not to take pleasure in Jones's pain", "it is wrong of you to want Jones to be injured", and even "you morally ought to take Smith's welfare into account in deciding what to do". If we are to take these utterances at face value, which I see no reason not to, and accept them as expressing moral truths just as much as do utterances like "you ought not to kill" and "it is wrong of you to steal", then it follows that people ought not to want or take pleasure in other people's pain and suffering, and that people ought to take the well-being of others into account in their deliberations. These moral requirements in conjunction with **Wrongdoing** explain the blameworthiness of Adam, Eve, and Ray in their respective situations.

6. A Radical Conclusion

I have argued that **Wrongdoing**, in conjunction with the true-blame-emotion conception of blameworthiness, can account for the blameworthiness of Adam, Eve, and Brad and, thus, is to

be preferred to **Wrong Behavior**. But don't the problems I raised for **Wrong Behavior** in section 3 similarly arise for **Wrongdoing**? There, I discussed two distinct problems for **Wrong Behavior**. One was that a person can be blameworthy even though she does not behave wrongly. The other was that a person can behave wrongly without being blameworthy. In this section I will examine how **Wrongdoing** fares with respect to these problems.

That there are situations in which people are blameworthy even though they do not behave wrongly will be a problem for **Wrongdoing** only if in those situations it is not the case that those people violate any moral requirements. But, as I shall now show, in such situations, though the people in question do not behave wrongly, and thus do not violate any moral requirements concerning behavior, they do, nonetheless, violate other moral requirements. Take, first, UNWITTING DEFENDER. In this scenario, Bob's killing of Sue is not a morally wrong action because, in doing so, he prevents a thousand innocent people from being villainously murdered by Sue. But, though he does not act wrongly, and, thus, does not violate a moral requirement not to kill Sue, he, nonetheless, does violate a moral requirement, namely, the moral requirement that he not be motivated to kill Sue merely for the sake of killing her. It is because Bob violates this moral requirement that he is blameworthy in UNWITTING DEFENDER. Had Bob known that Sue was about to kill a thousand people and been motivated to act by a desire to prevent her from doing so, then he would not violate this requirement and, thus, he would not be blameworthy for anything. But, as he is ignorant of these aspects of the situation, he acts on no such motivation and, thus, he does violate the moral requirement that he not be motivated to kill Sue merely for the sake of it.

Something similar is true with respect to TROLLEY. Recall that, in TROLLEY, James kills Chris by turning a trolley onto him. His turning of the trolley is morally permissible because had he not turned it, it would have barreled into and killed five innocent people. In turning the trolley and thereby killing Chris, James, however, is not motivated by a desire to save the five innocent people, but, rather, by a desire to kill Chris for his own profit. In this situation, James, though not violating a moral requirement not to turn the trolley, like Bob, violates a moral requirement concerning motivation. He violates the requirement that he not be motivated to kill Chris simply for his own profit. Had he been motivated to save the five instead of being motivated just to kill Chris, then he would not have been blameworthy for anything in TROLLEY. In both UNWITTING DEFENDER and TROLLEY, then, we see that each person, though he does not violate a moral

requirement concerning action, does violate a moral requirement concerning motivation. Cases like these, in which a person is blameworthy even though she does not act wrongly, thus aren't a problem for **Wrongdoing**.

Though cases like UNWITTING DEFENDER and TROLLEY do not refute **Wrongdoing** they may seem to raise the following problem: if in UNWITTING DEFENDER Bob only violates the moral requirement not to be motivated to kill Sue just for the sake of killing her, and in TROLLEY James only violates the moral requirement not to be motivated to kill Chris for his own financial gain, then is it not wrong to say that in UNWITTING DEFENDER Bob is blameworthy for killing Sue and that in TROLLEY James is blameworthy for killing Chris? Should it not be the case that Bob and James are only blameworthy for their motivations in these cases? On the true-blame-emotion conception of blameworthiness, a person is blameworthy for ϕ -ing only if it is appropriate for her to be blamed for ϕ -ing, and according to **Wrongdoing** a person is appropriately blamed for ϕ -ing iff in ϕ -ing she violates a moral requirement. But, if the only ϕ -ing of X's which is a violation of a moral requirement is X's having a certain motivation, then the only thing for which X is appropriately blamed is X's having that motivation. From this it should then follow that, in such a case, the only thing for which X is blameworthy is his having that motivation.

The foregoing is a problem only if in UNWITTING DEFENDER Bob is blameworthy for killing Sue and in TROLLEY James is blameworthy for killing Chris. But, I contend, it isn't clear that in these cases they are blameworthy for their killings. In cases like these, we are far more confident that the people in question are blameworthy, i.e., blameworthy for something, than we are about what, in particular, they are blameworthy for. There is, as a matter of fact, a bit of resistance to the thought that Bob and James are blameworthy *for killing their victims*. After all, we are inclined to think, they didn't perform a wrong action in doing so. Thus, I do not think that it is a datum that needs to be accommodated by a theory of blameworthiness that it yield that Bob and James are blameworthy for killing their victims. In sum, **Wrongdoing**, in conjunction with the true-blame-emotion conception of blameworthiness, does entail that Bob and James are not blameworthy for their actions, but this is not an objectionable entailment. Their being blameworthy for their morally wrong motivations is enough.

Unlike the first problem, however, the existence of blameless wrongdoing poses a much more significant problem for **Wrongdoing**. If it is the case that one can act wrongly but not be

blameworthy for anything at all, then **Wrongdoing** is in trouble. Take SUGAR. In this scenario, recall, Bill unwittingly poisons his neighbor, because, unbeknownst to him, Sally has replaced the sugar in his sugar bowl with arsenic. In this scenario it is clear that Bill is not blameworthy for killing his neighbor. If anyone is blameworthy for that, Sally is. What's more, it seems, Bill isn't blameworthy *for anything*. This is major trouble for **Wrongdoing** because it, in conjunction with the true-blame-emotion conception of blameworthiness, entails that if a person violates a moral requirement in ϕ -ing then she is blameworthy for ϕ -ing. Bill is a counterexample to **Wrongdoing** because he violates a moral requirement, namely, the moral requirement not to kill his neighbor, but he is not blameworthy for anything. Cases of blameless wrongdoing show that if the true-blame-emotion conception of blameworthiness is correct, then the content of a blame emotion felt toward X for ϕ -ing *cannot* just be that X violated some moral requirement. **Wrongdoing**, then, has to go.

Perhaps the blame emotions are responsive not to just any violation of a moral requirement, but, rather, to the violation of a particular class of moral requirements. If this is right, that would allow for there to be blameless wrongdoing in cases in which the person who commits the wrong does not violate one of the special class of moral requirements that the content of the blame emotions tracks. We have already seen that the content of the blame emotions must track violations of certain moral requirements that do not concern behavior. In particular, it must track the violation of moral requirements such as the requirement not to want or take pleasure in another person's suffering as well as the requirement to take the well-being of others into consideration in one's deliberations. In order to account for cases like UNWITTING DEFENDER and TROLLEY we saw that it must also track the violation of moral requirements that people not have certain malign motivations. The lesson of SUGAR seems to be that the moral requirement not to kill someone is *not* one of the moral requirements the violation of which is tracked by the content of the blame emotions. And so, if the requirement not to kill his neighbor is the only moral requirement that Bill violates in SUGAR, which seems plausible, then there is no blame emotion the content of which, were it felt toward Bill, would be true. Thus, as is correct, he is not blameworthy for anything in SUGAR. Bill is blameless in SUGAR, then, because though he violates a moral requirement, he does not violate one of the requirements to which the blame emotions are responsive.

If SUGAR does show that the blame emotions are not responsive to the moral requirement not to kill someone, then, it seems, a rather startling conclusion is just around the corner. As a case like SUGAR, in which someone, because of ignorance, is blameless for behaving wrongly, can, it seems, be constructed for any instance of wrong behavior, it follows from the true-blame-emotion conception of blameworthiness that moral requirements concerning behavior, in general, are *not* the kinds of moral requirements to which the blame emotions are responsive. This is certainly a surprising conclusion because, ordinarily, it seems that the things for which people are, in the main, blameworthy are instances of behavior.

Does this spell doom for the true-blame-emotion conception of blameworthiness? No, for the lessons learned in connection with UNWITTING DEFENDER and TROLLEY can be put to use here. In UNWITTING DEFENDER and TROLLEY it was shown that what the people in question were blameworthy for was not their actions, but the motivations from which they acted. Perhaps the only things people are ever blameworthy for when they perform actions are the motivations from which they perform them. If all of this is right, then, strictly speaking, no one is ever blameworthy for their actions. Rather, to the extent that people are blameworthy, in those cases in which we ordinarily think that they are blameworthy for their actions, they are actually blameworthy for the morally wrong motivations that those actions manifest. In most cases, however, wrong actions spring from wrong motivations. And so, ordinarily, our judgments about other people's blameworthiness for their wrong actions do track the actual blameworthiness facts. It's just that what people are actually blameworthy for are the motivations from which those actions spring, rather than the actions themselves. It is only in cases when wrong action comes apart from wrong motivation, as it does most often in cases of ignorance, like SUGAR and UNWITTING DEFENDER (though not only in cases of ignorance, as evidenced by TROLLEY), that we see that it is to wrong motivation and not to wrong action that blame is appropriately responsive.

In cases like SUGAR, in which a person is blameless for acting wrongly because of her ignorance, ignorance cuts the connection between wrong action and wrong motivation. If what blameworthiness tracks are things like the wrongness of a person's motivations, rather than that of her actions, then it is not at all surprising that ignorance can play the exculpating role that it often does. Ignorance is exculpating, on this theory, in just those cases where it leads someone

with a morally innocuous motivation to perform a morally prohibited action. This relatively neat explanation of how it is that ignorance is often exculpatory, then, is a mark in favor of the theory.

If it is a mark in favor of this theory that it has a neat explanation of how ignorance exculpates when it does, it is yet a further mark in its favor that that very explanation also explains why and in what cases ignorance fails to exculpate. Consider the following case:

ARAB BASHER: John intends to bash a Muslim but assaults an Arab Christian instead because he mistakenly, though blamelessly, believes that all Arabs are Muslims. If John had known that the person he bashed was not a Muslim, he would not have bashed him.¹⁸

ARAB BASHER is a case in which a person acts wrongly from blameless ignorance. In this respect, it is like SUGAR. Unlike Bill in SUGAR, however, in ARAB BASHER, John is blameworthy. The difference between the two cases is this. Whereas in SUGAR Bill's ignorance hinders his morally innocuous motivation's translating into a morally innocuous action, in ARAB BASHER, John's motivation is not morally innocuous to begin with. John's ignorance does sever the connection between his motivation and his action because, as a result of his ignorance, his motivation to bash a Muslim does not translate into his actually bashing one. But, as his motivation to bash a Muslim is itself a morally wrong motivation, John is not off the hook. As a matter of fact, John is as blameworthy in ARAB BASHER as he would have been had he successfully bashed a Muslim. According to the theory presently on offer, neither Bill, in SUGAR, nor John, in ARAB BASHER, is blameworthy for their actions. But whereas Bill is not blameworthy for being motivated to spoon sugar into his neighbor's tea, John is blameworthy for being motivated to bash a Muslim. Furthermore, this is explained by the fact that, though neither Bill nor John is bound by a moral requirement not to be motivated to spoon sugar into their neighbor's tea, both are under a moral requirement not to be motivated to bash a Muslim (or anyone for that matter).

Another consequence of this theory is that there is no such thing as consequential moral luck. Nagel (1976) has argued that there are a variety of different cases in which luck plays a determinative role in our moral lives. In the case of a type of luck he dubbed "consequential moral luck", he claimed that the degree to which a person is blameworthy for something can depend on luck with respect to the consequences of what she does. Now surely we do often blame people to different degrees depending on what the consequences of their actions are. But,

¹⁸ This is a version of an example taken from (Rosen 2002).

it seems to me, if two agents perform the same action with the same motivation and in the same circumstances, the two agents are equally blameworthy, regardless of the different consequences of their respective actions.¹⁹ For example, if Jones and Smith both independently shoot at Bloggs with the intent of killing him in order to inherit his fortune, Jones is no less blameworthy than is Smith even if Jones's bullet never makes it to Bloggs because it strikes a bird in mid-flight. Though Smith, and not Jones, kills Bloggs, and so he violates a much more serious behavioral moral requirement than does Jones—Jones violates only the behavioral requirement not to attempt to kill someone (and the requirement not to kill birds)—there is no level of indignation or resentment that it would be appropriate to feel toward Smith that it would not be appropriate to feel toward Jones. If this is right, then the one is no less blameworthy than the other. This fact is nicely accounted for by the theory presently under consideration because, as regards the moral wrong that the blame emotions track, i.e., the wrongness of motivations, etc., Jones and Smith violate the very same moral requirement—namely, the requirement not to be motivated to kill Bloggs in order to inherit his fortune.

(This is not to deny that there are certain emotions that it may be appropriate to feel toward Smith and his action that it would not be appropriate to feel toward Jones and his. It may very well be that the horror that it is appropriate to feel toward Smith and his murder of Bloggs would not be appropriate to feel toward Jones and his mere failed attempt. But, though there may be these other emotions, they are not the emotions constitutive of blame and, thus, they are no threat to my argument that, in this situation, Jones is no less blameworthy than is Smith.)

The current proposal—that the blame emotions are responsive to the violations of a set of moral requirements not concerning behavior—then, seems to offer a nice account both of the intuitive phenomenon of exculpation due to ignorance and of why Smith seems no less blameworthy than Jones in the above described scenario. But can more be said about the class of moral requirements the blame emotions track other than that it excludes the set of requirements concerning behavior? Is the relevant class merely the complement of that set? No. There surely are moral requirements that are not requirements concerning action, but which are also ones that

¹⁹ Here I am in partial agreement with Michael Zimmerman. Zimmerman (2002) has argued that there is no moral luck of any sort. I do think that there are some types of moral luck. For example, it may be a matter of luck which community one is born into and, thus, a matter of luck what kind of character one develops as a result of growing up in that community. In such a case it may very well then be a matter of luck whether, in virtue of having the character that one has, one violates one of the norms that blame tracks. In this way it can be a matter of luck whether someone is blameworthy for something. So, I do think that there is such a thing as moral luck, I just don't think that there is any of the consequential variety.

can be blamelessly violated. Though there is a moral requirement that I not be on your land (and not simply that I not venture onto it), if out of blameless ignorance I end up on your land, I am not blameworthy for anything. The set of moral requirements to which the blame emotions are responsive, then, is not just the set of moral requirements other than those concerning behavior. What, then, unifies them as a class? As we have seen, this class includes requirements concerning which motivations to have and not to have, requirements not to want or enjoy other people's suffering, as well as requirements to treat other people's interests as significant in one's deliberations. Though this group of moral requirements may seem a bit motley, what unifies them is that they are requirements that constrain the psychic elements that compose one's character. To give this class of moral requirements a name, I dub them "requirements of character". The theory of the content of the blame emotions for which I have to this point been arguing is:

Character: the content of a blame emotion felt toward X in response to X's ϕ -ing is that, in ϕ -ing, X has violated a moral requirement of character.

Moral requirements of character constrain the elements of our character in such a way that we display a sufficient amount of respect toward others. Our value as persons places on others a moral requirement to accord us the respect that that value demands. Insofar as we bear malign intentions toward others, want or hope for their suffering, or fail to take their well-being into account, we fail to accord them the respect that their intrinsic moral worth requires of us.

In addition to accommodating much of the intuitive blameworthiness data, **Character** also intuitively fits with the thought that blame is a response to persons and not to actions. Intuitively, we do blame people for actions. But the actions are important only insofar as they reflect negatively on the person who performs them. This is preserved by the theory I am offering, for according to that theory what people are really blameworthy for are those features of themselves that pretheoretically their blameworthy actions reflect. On the theory I am propounding, what people are, strictly speaking, blameworthy for when they act are the motivations from which they act. But, in a loose sense, they can be said to be blameworthy for their actions to the degree that those actions are the natural and direct manifestations of those motivations. In this way, though the official line of the theory is that no one is blameworthy for their actions, the theory can accommodate the pretheoretic thought that people are blameworthy for their actions.

7. Conclusion

The theory of blameworthiness I have sketched comprises three crucial elements. (1) To be blameworthy for ϕ -ing is to be such that it would be appropriate for the blame emotions to be felt toward one on account of one's ϕ -ing. (2) The blame emotions are appropriately felt just in case the propositional content of those emotions is true. And (3) the content of a blame emotion felt toward a person for ϕ -ing is that in ϕ -ing that person has violated a moral requirement of character. This theory accommodates most of our intuitive judgments about blameworthiness at the price of entailing that people are not, strictly speaking, ever blameworthy for their actions. Rather, on this theory, in those cases in which, intuitively, they seem blameworthy for their actions, they are, instead really blameworthy for the motivations from which those actions spring. On balance, however, this is a small price to pay. For in recompense, the theory offers a unified account of the multitude of things for which a person can be blameworthy as well as offering a simple explanation of why and when ignorance is and is not exculpatory.²⁰

²⁰ I would like to thank Yuval Avnur, Anne Barnhill, Paul Boghossian, Jessica Boyd, Alex Guerrero, Matthew Kotzen, Peter Kung, Michael Otsuka, Ryan Preston, Erica Roedder, Karl Schafer, Josh Schechter, Brad Skow, Declan Smithies, Masahiro Yamada, and participants in the NYU dissertation seminar for helpful comments on earlier drafts of this paper. Extra special thanks go to Cian Dorr, Elizabeth Harman, Thomas Nagel, and Gideon Rosen for numerous discussions of the material in this paper as well as extensive comments on many successive drafts.

REFERENCES

- Audi, Robert (1991), "Responsible Action and Virtuous Character", *Ethics* 101: 304-321.
- D'Arms, Justin and Daniel Jacobson (2000), "The Moralistic Fallacy: On the 'Appropriateness' of Emotions", *Philosophy and Phenomenological Research* 61: 65-90.
- Deigh, John (1994), "Cognitivism in the Theory of Emotions", *Ethics* 104: 824-854.
- Hieronymi, Pamela (2004), "The Force and Fairness of Blame", *Philosophical Perspectives XVIII: Ethics* 115-149.
- Kamm, F. M. (2001), "Toward the Essence of Nonconsequentialism", in A. Byrne, R. Stalnaker, and R. Wedgwood, eds., *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson* (Cambridge Mass.: MIT Press).
- Nagel, Thomas (1976), "Moral Luck". Reprinted in Nagel (1979).
- Nagel, Thomas (1979), *Mortal Questions* (Cambridge: Cambridge University Press).
- Roberts, Robert C. (1988), "What an Emotion is: A Sketch", *The Philosophical Review* 97: 183-209.
- Rosen, Gideon (2002), "Culpability and Ignorance", *Proceedings of the Aristotelian Society* 103: 61-84.
- Smith, Angela M. (2005), "Responsibility for Attitudes: Activity and Passivity in Mental Life", *Ethics* 115: 236-271.
- Strawson, P. F. (1962), "Freedom and Resentment", *Proceedings of the British Academy*; in Watson, (2003) 72-93.
- Wallace, R. Jay (1994), *Responsibility and the Moral Sentiments* (Cambridge Mass.: Harvard University Press).
- Watson, Gary (2003), *Free Will* (Oxford: Oxford University Press).
- Zimmerman, Michael J. (2002), "Taking Luck Seriously", *The Journal of Philosophy* 99: 553-576.