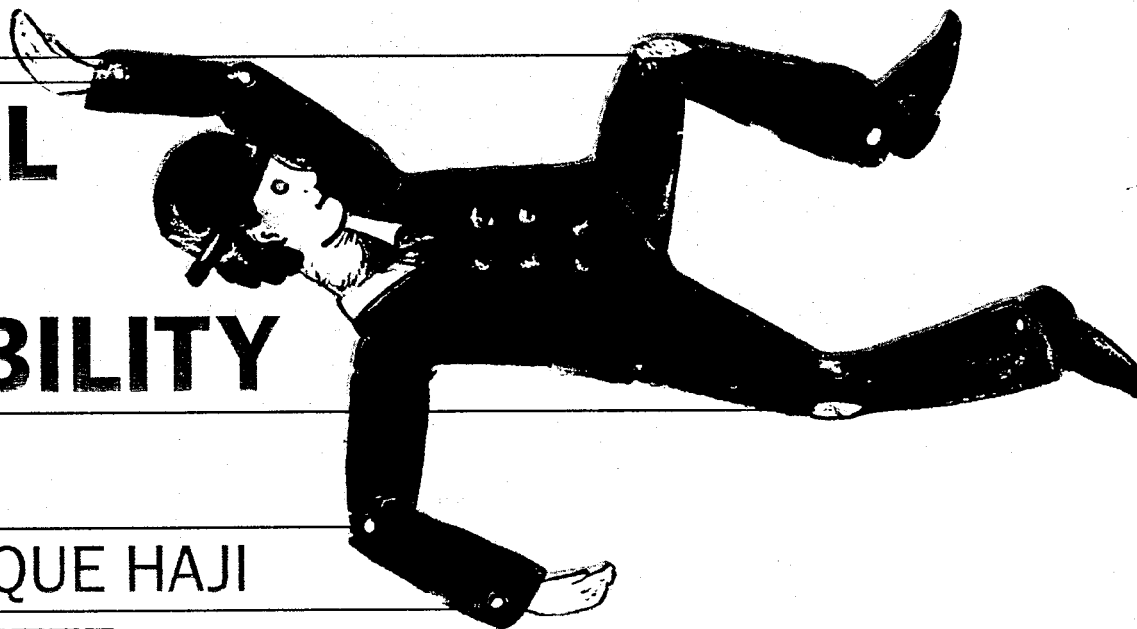


**FREE WILL  
AND MORAL  
RESPONSIBILITY**

EDITED BY ISHTIYAQUE HAJI  
AND JUSTIN CAQUETTE



proposal is false. What I should have said, roughly, is that one need not be praiseworthy for performing something akin to a supererogatory action. The idea is that, to the extent this is possible, imagine an action that has all the features of a supererogatory one save (roughly) that the agent performs it “from” base motives. Then we may have a supererogatory-like action for which an agent (such as Marion in our example) is not praiseworthy. McNamara also discusses various interesting cases in which an agent can be blameworthy (and praiseworthy) for doing something that was either precluded by doing the minimum or precluded by doing the maximum in, e.g., 2011a; 2011b; 2008.

<sup>9</sup> On owed favors, see, for e.g., Driver 1992, p. 289.

<sup>10</sup> See, for instance, McNamara 2011a, p. 231. See, also, e.g., Chisholm 1963; Mellema 1991; and Driver 1992.

<sup>11</sup> Again, previously, not having an analysis of the suberogatory, I suggested that it is possible to do something that is suberogatory without being blameworthy for doing this thing (2002, pp. 178-79). If McNamara’s analysis is correct, this is false. What I should have said, roughly, is that one need not be blameworthy for performing something akin to a suberogatory action.

<sup>12</sup> Jay Wallace (1994) also ties blameworthiness to quality of will. He proposes that one is blameworthy for performing an action only if in so doing one violated a moral obligation, and one’s moral obligations concern one’s quality of will as expressed in action.

<sup>13</sup> We have already registered that one may discharge an obligation in various ways; some of these ways may be impermissible.

<sup>14</sup> Here, as elsewhere in the paper, I do not address what may be other problems for different varieties of semicompatibilism.

## CHAPTER EIGHT

### FREEDOM AND FORGIVENESS

DANA NELKIN

“To understand is to forgive”

—proverb

“Forgiveness is not about the person who transgresses you, it is about you.”

—Dr. Phil

“Always forgive your enemies, nothing annoys them so much.”

—Oscar Wilde

#### 1. Introduction

In his influential and important paper, “Freedom and Resentment,” Peter Strawson introduced the term of art, “reactive attitudes” a term that has gone on to make frequent and central appearances in subsequent discussions of free will and moral responsibility. He described the reactive attitudes as “the non-detached attitudes and reactions of people directly involved in transactions with each other; of the attitudes and reactions of offended parties and beneficiaries; of such things as gratitude, resentment, forgiveness, love, and hurt feelings” (2003 p.75). About forgiveness in particular, Strawson remarked that it is “a rather unfashionable subject in moral philosophy at present.” This was probably true at the time, but I do not believe it would be if written today. And not only is the subject receiving more attention by philosophers, it is at the same time receiving a great deal of attention by psychologists, medical professionals, and self-help practitioners, among others.

The latter groups are naturally interested largely in the psychological, medical, and spiritual effects of various mental phenomena, and one result is that there can be a lack of clarity and consistency about what the phenomenon in question is, and how it is distinguished from other

phenomena that might be similar in some of their effects. For example, forgiveness is not always clearly distinguished from related phenomena that seem intuitively different on reflection, such as “moving on” or “forgetting” or “distraction.” All of these phenomena are associated with the elimination, or at least mitigation, of negative feelings such as resentment, for example. But as it turns out, there may be important differences between all of these phenomena, even in their psychological and physical effects.<sup>1</sup> And so a good philosophical account of the nature of forgiveness would seem to have important potential practical applications, in addition to its importance for ethics.

In fact, the confusion about the nature of forgiveness and how to distinguish it from other phenomena turns out to be quite understandable. Philosophers have themselves struggled with a variety of related puzzles and even so-called “paradoxes” of forgiveness.<sup>2</sup> Here is one such puzzle: forgiveness seems to be distinct from mere forgetting or not thinking about the relevant offense, and a natural suggestion is that forgiveness involves some change in attitude toward the offender and offense. As the proverb goes, in forgiveness one achieves some sort of understanding one didn’t have before. But what sort of understanding? If one comes to a different and deeper understanding of the causes of the offense than one previously had, for example, this could lead us to a different phenomenon altogether—namely, excuse. Imagine that you resent your father for having been overly critical of some of your decisions as a teenager, resulting in an overly cautious approach to life later. If now, perhaps by seeing things from his point of view, you come to see how influenced he was by aspects of his own early environment and by his own pressing fears, you might understand things better, but might also find yourself withdrawing your initial judgment of responsibility, and so having nothing to forgive. So in forgiving—if it is really a distinct phenomenon from excuse—it seems we must continue in our initial judgment that the offender is blameworthy for committing the offense, while at the same time change in some way in addition to (or other than) having one’s feelings of resentment fade. The challenge is to navigate between excuse on the one side and, on the other, phenomena like forgetting that do not seem to require any new action or attitude directed at the offender or offense.

Meeting this challenge and getting clear on the nature of forgiveness is important on its own terms. But it is also important in ways that are related to the debate over whether we are free and responsible agents. First, it forces us to think about the relationship between forgiveness and what philosophers at least have thought to be presupposed by it—the free

and responsible agency of the offender. Powerful skeptical arguments challenge the claim that we are in fact free and responsible; if they are correct, then all of our attitudes that presuppose that we are free agents rest on an illusion and are unjustified.<sup>3</sup> If forgiveness must be understood as having this presupposition, then the appropriateness of forgiveness itself is at stake in the debate between the skeptics and non-skeptics about freedom.<sup>4</sup>

Second, the challenge mirrors one aspect of the debate about the nature of freedom itself. The debate between compatibilists and incompatibilists about freedom is whether any agent can be free in a determined world. In focusing on determinism, we understand that everything is caused by prior events, in turn caused by prior events, and so on, in accordance with natural laws. If we see a particular person’s transgression as caused in this way, we understand it better in a sense, but many are inclined at the same time to withdraw attributions of responsibility in such circumstances. If we can account for forgiveness as involving understanding without withdrawing such attributions, we might be able to obtain some insight into this more abstract debate.

In this paper, I will make a start on this project by proposing an account of an important set of central cases of forgiveness. In the next section, I will briefly set out *desiderata* that a good account of forgiveness should satisfy, and in section III, I will canvas some of the best accounts available. Even where they fall short, I believe that we can learn a great deal from them. In section IV, I will propose an account that incorporates the insights gained, and in section V, I will briefly return to the parallel between forgiveness and the debates about freedom and responsibility.

## 2. Desiderata

Although it would be nice to have a neat set of necessary and sufficient conditions that cover every instance of what people on reflection would call forgiveness, this is unlikely to be possible. And for our purposes it would be well worthwhile to have an account that captures central cases of forgiveness and that at the same time shows why forgiveness is distinct from the variety of phenomena mentioned earlier. There may be some gray areas where it is not clear whether something is a genuine case of forgiveness, and ideally our account will be able to explain why it falls into a gray area. Thus, a set of conditions that does these things, explaining what is characteristic and essential to central cases, is my aim here.<sup>5</sup>

It is important at the same time not to restrict central cases too narrowly, or we may mistakenly take features to be essential that are not. For example, although forgiveness very often follows apology and repentance, there seem to be many cases that do not include either of these things. If we limit ourselves to cases of apology, we risk imposing particular requirements on the offender, for example, that do not apply to forgiveness in general.

As the earlier discussion brought out, if we are to distinguish forgiveness from excuse, we must preserve the idea that the forgiver continues in her attribution of responsibility to the offender for the offense, and at the same time explain the change that constitutes forgiveness in a different way.

Finally, a large part of the literature on forgiveness is focused on a variety of interesting normative questions, and it would be a virtue of an account of the nature of forgiveness that it points the way to how these might be answered. For example, is forgiveness ever required, is it ever impermissible, is it ever supererogatory, and if so, under what conditions? An account of the nature of forgiveness need not answer these all by itself, but it should, combined with more general moral principles, point to answers.

### 3. Three Recent Accounts of Forgiveness<sup>6</sup>

#### Overcoming of Resentment on Moral Grounds

Most philosophers writing about forgiveness take it that a necessary, if not sufficient, condition for forgiveness is the overcoming of resentment. (Many cite Bishop Butler as an earlier source of this idea, although what he actually says is that forgiveness is the overcoming of "excessive" resentment.<sup>7</sup>) In what is perhaps the most influential contemporary account of forgiveness, Jeffrie Murphy makes the overcoming of resentment the centerpiece of his account, but takes it to be insufficient on its own. Forgetting is also a way of losing resentment, he points out, as is behavior-modification therapy, but neither is the same thing as forgiveness.<sup>8</sup> We must add a condition, Murphy argues, and that is that the forswearing of resentment must be done "on moral grounds" in order to count as forgiveness. Murphy is explicit that the question of what forgiveness is and what *justified* forgiveness is cannot be sharply distinguished. But they can be distinguished to an extent, for there are better and worse moral grounds, or reasons. What is essential for forgiveness is that it be for the right *kind* of reasons (even if their force may be weak in the context of

countervailing reasons). What count as reasons of the right kind? Murphy lists five: the offender repented or had a change of heart; he meant well; he has suffered enough; he has undergone humiliation (which might include a kind of apology); and for old times' sake (Murphy and Hampton 1988, p. 24). Although they appear a diverse lot, Murphy argues that they fall under a unifying umbrella: they are all ways of separating act and agent. Thus, on Murphy's view, forgiveness is the forswearing of resentment for one of these reasons that instantiates the separation of act and agent.

This account has many virtues. It is simple and elegant, and is presented with an eye to distinguishing it from the variety of phenomena with which we began. At the same time, it is vulnerable to certain objections. Consider, first, the claim that the unifying category of separating act from agent captures all of the specific moral grounds identified by Murphy. The fact that an agent has suffered already, for example, does not seem to bear in any way on the relationship between the agent and the particular offense in question. The well-known case of Robert Alton Harris who committed atrocious crimes, but who had suffered horribly from the treatment from his abusive parents supports this point.<sup>9</sup> Harris' own suffering might provide some reason for forgiveness, but it does nothing to separate his murders from himself. Perhaps Murphy could simply stand by the claim that the grounds must all be "moral", while withdrawing the more specific claim about what unites them. But this cannot be quite right either, for, as pointed out by Jean Hampton, there may be other moral grounds available, such as the fact that forgiving will help keep peace within the family, that seem not to be of the sort Murphy himself has in mind.<sup>10</sup> The view might be modified, it seems to me, to take this point into account: the grounds in question must be moral *and* directed toward the offender (as opposed to, say, one's family, or oneself).

Yet even if this issue of how to understand the "moral grounds" could be resolved, three serious challenges remain. The first is that one does not actually need the restriction to moral grounds to distinguish forgiveness and forgetting. As Murphy also seems to be aware, forgiveness is active, whereas forgetting is something that happens to you. The existence of grounds in the former case is not needed to make this distinction. Second, and perhaps more importantly, we have not been given sufficient reason for thinking that grounds of any kind (good or bad) are needed. Perhaps in typical cases, we adopt reasons when we forswear our resentment. But I do not think we need to, even when we forgive. Just as we may act for bad reasons, we may act for none at all. Take an analogy: belief. Typically, we believe on the basis of reasons; but sometimes our belief is irrational and even sometimes motivated. We believe *against* the evidence

we have on occasion. Similarly, it seems possible to forswear resentment not because we have adopted good (or bad) reasons that support our doing so. Rather, doing so might satisfy a desire or motivation that we have. It is not clear why, once we allow that we can forgive for bad reasons, we cannot forgive on the basis of no reasons. This is a point to which I will return.

Third, as even Butler himself seems to allow, forgiveness appears to be consistent with at least some feelings of resentment. I might have forgiven an old teacher for her harshness even though I find myself with occasional feelings of resentment.

In sum, Murphy's view represents a very widespread view of forgiveness, and offers clear ways of addressing various *desiderata*. At the same time, his view raises important questions about the role of reasons in forgiveness, as well as the necessity of forswearing resentment. I now turn to a different view that nevertheless shares with this one some key points.

### Undermining of the Perception of Threat

Pamela Hieronymi lays out very clearly a puzzle similar in key ways to the one with which we began. Two fixed points must be preserved by the correct account on her view. First, an account of forgiveness must, in her words, be "articulate," in order to distinguish it from phenomena like forgetting or distraction. (In other words, it must include a change in judgment.) The mere fading away of resentment does not constitute forgiveness; on her view, as on Murphy's, forgiveness is the overcoming of resentment, but that overcoming must be sensitive to reasons in the form of a change in judgment. Second, the correct account must be "uncompromising" in not requiring the change in judgment to be a backing away from one's attribution of responsibility and judgment that one shouldn't be treated in the way one was.

Focusing on cases in which apology and repentance are tendered, Hieronymi offers such an articulate and uncompromising account. She begins by understanding resentment as a "protest against a past action that persists as a present threat" (2001, p. 546). The past action poses a threat because a past action that wronged you in a significant way and that has not been atoned for, repudiated, or apologized for, makes a continuing claim that "you can be treated in this way and that such treatment is acceptable." Resentment stands as a denial of this claim, and a protest of the threat is carried to either the public understanding or one's own. One's judgment that there is a continuing threat makes resentment appropriate,

and constitutes a reasonable inference from one's judgment that one has been significantly wronged by a responsible agent. Resentment need not entail a desire that the offender suffer for it, but can include a desire for apology. Given this understanding of resentment, apology can then bring about a change in judgment by cancelling the earlier judgment that the past action continues to make the threatening claim. At the same time, it does *not* cancel the original judgments of responsible and unacceptable wrongdoing.

While Hieronymi explicitly focuses most of her attention on the case of apology, she is also concerned to extend the account to other cases. Perhaps, she writes, forgiveness of the unrepentant is possible "in cases in which the one offended receives strong community support." If one is assured that the wrong will be acknowledged even without her resentment, that the wrongdoer will continue to be expected to do otherwise, and if she herself is duly respected by the community, then, Hieronymi concludes, perhaps the judgment that there is not a continuing threat may reasonably be made, even without an apology (2001, pp. 552-53).

This account also has a welcome simplicity, and neatly satisfies the dual requirements of articulateness and uncompromisingness. But it also falls short of capturing essential aspects of forgiveness in what I believe are instructive ways. First, as Hieronymi herself is aware, one might raise a number of questions about what it means for an action to express a continuing claim. Relatedly, it isn't clear that resentment has as its object a claim, even one that is revealing of the offender's evaluations. I will set aside such questions for now, however, and turn to two different concerns.

Imagine the situation of two parents whose child has been murdered. There might be apology on the part of the offender as well as genuine repentance, and all the community support one could ask for. Suppose the parents recognize all of this, and judge that the past action no longer makes the kind of threatening claim Hieronymi has described. It does not seem that they need have thereby forgiven. Perhaps Hieronymi could agree—perhaps we could read her view as claiming only that the parents would no longer *have reason to* resent, and would have reason to overcome such feelings. In fact, she articulates something like this view in anticipating an objection from those who have thought forgiveness is always "elective" or optional when she writes: "an articulate account articulates the reasons which *call for (rationally require)* forgiveness." (2001, p. 551, emphasis mine).<sup>11</sup> But even understood as an account of the conditions under which forgiveness is rationally required, it seems too easy to satisfy. In the case of the parents, I believe that we could imagine there being no reason to criticize them for their failure to forgive, whether

on moral or rational grounds. This suggests that the account is missing something important about what forgiveness is.

Now one might reply that it is not the parents' place to forgive in the first place, since they were not the targets of the offense. (Some have reserved 'indignation' as a third-party analogue for resentment; and have spoken of a kind of third-party analogue of forgiveness rather than forgiveness proper for such cases.) This may be a case in the gray area, since the action could be seen as an offense against not only the child but against those who loved and valued the child. But purer cases are available, as well. One might have been the victim of a brutal assault or rape, and, again, received genuine repentance from the offender as well as vast community support so that one is assured of respect in one's own eyes as well as one's community's. Yet, here, too, it seems that one needn't be making a mistake, either moral or rational, by not forgiving.

This objection targets any claimed sufficiency for the account. But we might also question whether the conditions proposed are necessary. Understood in the way that Hieronymi does, resentment is a kind of protest. Similarly, as Gary Watson sees the reactive attitudes, they embody a "moral demand" to be treated with respect.<sup>12</sup> It seems possible in this sense to continue to resent while also having forgiven, just as it seems possible to protest the wrong (and any continuing claim) after having done so.

Because of the richness of Hieronymi's account, even where I have argued that it falls short, it is possible to take certain lessons from the challenges it faces. In one way, more is required to forgive—more than simply a change in judgment about a continuing threat. In another, less may be required—resentment, like blame more generally, might be a multi-faceted thing. And so resentment—or at least certain aspects of it—might be perfectly consistent with forgiveness.

### Overcoming of Ill Will and Adoption of Good Will

In a series of papers and a book, Garrard and McNaughton develop a nuanced account of forgiveness as the overcoming, not of resentment *per se*, but of *hostile* feelings or feelings of ill will that may be a part of resentment, together with some minimal adoption of good will.<sup>13</sup> Ill will is understood as "wishing harm to someone, relishing the pain and discomfort that he suffers" (2010, p. 23). Notably, purging oneself of such feelings is consistent with the maintenance of a number of other negative feelings, including disappointment, and even indignation. Indignation in turn is understood as a kind of moral outrage at a moral wrong and even

perhaps as including the desire that the breach be both rectified and punished.

At the same time, it is important to note an important consequence, on their view, of understanding forgiveness as constituted by the overcoming of feelings of ill will. They argue that it follows that forgiveness "involves not requiring either apology or penance" (2003, p. 47). They derive the conclusion from the claim that apology and penance amount to a kind of humbling before one, and the claim that relishing the lowered standing these entail would undermine the idea that one no longer feels ill will. Filling in here, it seems that they take apology and penance to necessarily involve a kind of suffering or harm, and so requiring either one would suggest feelings of ill will after all.

Purging of ill will and its consequences are not the whole story, however. For such purging or overcoming is also consistent with disdain and contempt, and with thinking that the offender is not even worthy of consideration. If these latter feelings mean that one has no good will at all for the person, then one cannot have really forgiven. One must have at least some good will, as well.

Garrard and McNaughton also provide direct answers to the normative questions with which we began. They argue that it is always permissible to overcome ill will and so to forgive. The reason is that there are always good reasons to do so, most notably embodied in a sense of human solidarity. At the same time, forgiveness for bad reasons is not admirable. Thus, like Murphy, they can distinguish good and bad reasons for forgiveness, a feature of their account that seems true to the phenomenon.

In sum, then, on their account, forgiveness is the overcoming of specific feelings of ill will and the possession of at least some good will. There is always reason to forgive, and it is always admirable if done for the available good reasons.

The account and arguments contain many important insights. The idea that resentment and blame are complex phenomena appears to be crucial to addressing the puzzles about forgiveness. For, there is a way in which we continue to blame even after having forgiven (least controversially, for example, in judging that the person was responsible for wrongdoing), and yet ways in which we do not (for example, demanding apologies or forms of penance). This seems right, and dovetails nicely with a newly revived attention to the "ethics of blame" itself.<sup>14</sup>

And yet, I do not believe that the case has been made that the key to forgiveness turns specifically on feelings of ill will, or a desire for suffering or harm, as the account claims. It seems possible to imagine cases in which one forgives an offender for a particular offense without

ridding oneself of feelings of ill will. One might, for example, have forgiven someone for stealing to support a drug habit, but not for refusing to attend a rehabilitative program. At the same time, one might feel general ill will toward the offender. Or one might have forgiven for a recent transgression but not for an earlier one. These seem to constitute counterexamples to the claim that the rejection of ill will is even necessary for forgiveness. In response, the view might be clarified to account for this possibility by allowing feelings of ill will, so long as the feelings no longer draw any support from the offense that is allegedly forgiven. Ill will might remain, but not on account of that particular offense. This seems conceptually coherent, but it also seems psychologically difficult for people to distinguish the sources of their own feelings of ill will. It might be that it is simply hard for anyone—including forgivers—to know when they have forgiven and for what. But since we do often claim to know these things, it would be an advantage if an account could easily explain this.

Still, there is an even stronger kind of counterexample: consider people who have no such feelings to be overcome in the first place. It is quite conceivable that someone in such a situation forgives. It might help to see this if we consider that a lack of desire for suffering, as Garrard and McNaughton point out, is consistent with feelings of outrage and the making of various kinds of demands. If forgiveness requires the overcoming of ill will in particular, then even those who feel indignation and forswear it, without also desiring the offender's suffering, will not count as having forgiven.<sup>15</sup>

Finally, we might ask about the role that Garrard and McNaughton give to demands for apology and penance. While I fully agree that the victim's asking the offender for apology or penance would undermine our judgment that the victim had already forgiven, the reasoning offered by them for this conclusion is questionable.<sup>16</sup> On their view, a demand for apology would be undermining of a judgment of forgiveness precisely because it would signal some residual feeling of ill will. But this claim seems to rest in turn on the following implicit assumptions: (i) any candidate forgiver believes that both apology and penance entail suffering, and that (ii) the demand for apology and penance in turn entail a kind of "relishing" or desire for suffering. I do not believe that either of these assumptions is correct. (i) is questionable because it would at the very least take an argument that the mere making of a sincere apology entails suffering.<sup>17</sup> On its surface, it does not seem to require suffering. Apologizing might come as a relief or something welcome to the offender. Because it actually appears to be false that apology must involve suffering,

it is implausible to attribute the belief that it does to all forgivers. (ii) is also problematic. As Garrard and McNaughton argue elsewhere in the context of punishment, one might demand punishment, punishment might essentially involve suffering, and yet one need not necessarily desire the suffering for its own sake. Similar reasoning could be employed if we substitute "apology" for "punishment."

Thus, while I think Garrard and McNaughton are correct that forgiveness precludes a further demand for apology or penance, this claim is not in fact entailed by their canonical account of forgiveness, which is given entirely in terms of the forswearing of ill will and adoption of good will. This suggests that we look further for what it is that really explains the idea that forgiveness is not compatible with a continuing demand for either apology or penance.

#### 4. Developing a Debt Release Model

The most natural solution to this particular problem is to say that forgiveness is *constituted* (at least in part) by a special kind of release from a special kind of obligation the offender has to the victim. In typical cases, the obligation might be fulfilled by apology, sincere remorse, penance or related phenomena. In forgiving, one ceases to hold the offense against the offender, and this in turn means releasing them from a special kind of personal obligation incurred as the result of committing the wrong against one.

As we are often reminded, "forgiveness" has many meanings. Forgiving a debt, for example, is different from forgiving a person for an offense. But I think it is no coincidence that these two phenomena share the same name. Understanding interpersonal forgiveness for an offense on the model of a special kind of release from a special kind of debt or obligation turns out to have many attractions. First, if forgiving is conceived as a kind of personal release, we can understand immediately how it is different from forgetting or not thinking about the offense. The former, but not the latter, requires a kind of activity, a special kind of intention formation. This does not mean that forgiving requires a conscious decision, but it does require an intentional process. (Just as we can act without having first consciously decided to, so we can forgive.) Second, forgiveness is distinct from excuse, because a release from a personal obligation has no implications for a change in attribution of responsibility for the act, or so I will argue. Much will need to be explained, of course, including why forgiveness has been associated with the mitigation or elimination of resentment, how it is related to

understanding on the part of the victim, why it has been thought to be virtuous to forgive (or be a forgiving person), and so on.<sup>18</sup>

The model has many precedents. Some Christian models of God's forgiveness for human sins, for example, take God's forgiveness to be the cancellation of a debt we owe to God as our creator.<sup>19</sup> I, like the other authors whose work I have discussed, take myself to be trying to account for a secular notion of forgiveness. But it is important to note that this is a clear and influential case in which a kind of "interpersonal" forgiveness for an offense is conceived on a kind of debt-forgiveness model. The idea that the two notions of forgiveness are related is also suggested in contemporary and explicitly secular work—including that of Garrard and McNaughton who, as we have seen, believe that the release from apology or penance is a part of the correct account of forgiveness.<sup>20</sup> P. Twambley (1976) quite explicitly adopts such a model, sketching a parallel to mercy: "By offending you, a man, as it were, incurs a debt (hence we talk of owing recompense, reparation, and apology). You are within your rights to resent his action. In forgiving him, you relinquish that right, you readjust your relationship to one of equality" (p. 89).<sup>21</sup> And although Swinburne (1989) offers an account of forgiveness within a larger religious context, the interpersonal aspects can be understood on their own as a kind of debt release model, where the debt incurred in wrongdoing is owed to the victim (and not only to God).

Despite these precedents, however, many of the most influential models recently offered by philosophers seem to focus primarily on the elimination (or mitigation) of resentment or ill will or other negative emotions. The ceasing to "hold it against" the offender, and the release from debt, seem to be treated, when they are treated, as consequences of the mitigation of these negative emotions rather than as constituting forgiveness, as in Garrard and McNaughton's view.

In what follows, I will begin to develop a more specific version of the debt release model of forgiveness, and explain how it promises to address the *desiderata* with which we began. I will then turn to some objections to debt release models, and show how a particular version of the model can satisfy them.

People owe each other a variety of things for all kinds of reasons. We incur obligations all the time on the basis of voluntary actions for which we are not blameworthy. But when we wrongfully and culpably harm others, we incur at least two sorts of obligations: the obligation to make restitution for the loss or harm suffered (if we stole a bicycle, then we owe a bicycle or as close to the equivalent as we can provide), and the obligation to somehow make up for or in some way address the wrong

itself. The content of this second kind of obligation as I've characterized it is vague, and one might think we could and should be more precise by saying that one has the obligation to apologize, repent, do penance and otherwise "make up for it", possibly by giving extra of one's time or possessions in charitable ways.

I believe that this would be an acceptable emendation if we think of apologizing, repenting, doing penance or otherwise making up for the wrong as a disjunction, since one might be able to fulfill one's obligation in any one of multiple ways. But an important question arises here about whether the *extent* of one's obligation is a simple function of the severity and type of offense, or whether it can also vary depending on the victim's own view of what would be required to "make up for" the wrong, if anything.<sup>22</sup> In fact, there is a challenge for the view to provide an answer to this question.

It will help to present the challenge in the form of a dilemma. First, if we read the content of the obligation as simply being a function of the severity of the offense, then it seems it should be possible for the offender to say, "I've done everything I owed; now I am no longer under an obligation." In that case, there would be nothing for forgiveness to do. But even if we think (as I do) that there are occasions when we really *ought* to forgive, it still seems elective in one important sense, namely, that you don't lose *the option* of doing so or not simply because the offender has completed certain important tasks.

On the other hand, if we read the content of the obligation as something set by the victim, then we retain the electivity of forgiveness, but we also seem to risk losing the idea that there are circumstances in which one really ought to forgive. If the victim can move the goalposts, so to speak, then on what grounds could we ever object to the victim?

This seems to be a challenge for the debt release model.<sup>23</sup> I will sketch replies to the arguments against each horn of the dilemma, but in the end, tentatively conclude that the obligation is not set by the victim.

One might try to defend the view that the obligation can be set by the victim as follows: there is nothing problematic in thinking that what offenders are required to do is subject to the luck of the draw; they took the risk, and are now beholden to fulfill an obligation set by their victims. And even if the victim sets the obligation, we need not conclude that the victim has acted well in setting an unreasonable obligation. Finally, it remains the case on this view that a variety of different kinds of reasons—different from meeting the terms of the obligation—might support release from the obligation. In fact, it is a crucial part of the very idea of the debt



release model that the forgiver can release the offender from an obligation *not fulfilled*.<sup>24</sup>

Still, it seems possible that a victim reflects that the offender has done all he has been asked and can reasonably be asked to do, and more, and yet not forgive. This seems true whether the obligation is set by the victim or not. But if this is the case, the model must explain how the victim could still be in a position to choose to release the offender from the special obligation even after the offender has fulfilled its terms. I think that there are two ways the debt-release model can answer this. One option is to combine the debt release model with something like the overcoming of resentment model as a separate condition. For reasons given earlier, concerning the possibility of some continued feelings of resentment after forgiving, I am more tempted by a second kind of response.<sup>25</sup> It is to understand the fulfillment of the terms of the obligation as not entailing the actual release. (Analogously, one may have served his sentence to the minute, but not yet have been officially released.) There may be then either a kind of irrationality or a kind of moral failing in such a situation: one judges that an obligation has been fulfilled, but one refuses to release the obligated. This seems to be the kind of case in which one ought to forgive. There may be understandable and mitigating reasons why a victim does not, and we may not blame her for her failure, but it could still be true that she ought to do so. The upshot of this reply is that obligations are set by the circumstances and severity of the offense, but that even when the offender has met his obligation, he may not yet be forgiven. However, at this point, *if* the forgiver knows him to have met the conditions of the obligation, there is irrationality in refusing to release him.

So far, then, we have a sketch of how to characterize the debt incurred when one culpably commits a serious wrong, with options as to how to fill in the details. I have made tentative suggestions as to which options to pursue, but do not yet wish to close the door on the others. In continuing to fill out the model, it is important to ask whether there must be a change of judgment in order to forgive. A change of judgment plays a large role in many accounts of forgiveness (as in Hieronymi's account), and though we have already distinguished forgiveness from other phenomena like forgetting and distraction without it, it does seem intuitively as though some different way of looking at things is required.<sup>26</sup>

In identifying what constitutes the change of judgment required, I believe we should be guided by two sorts of constraints. The first is that the change of judgment should be related in an appropriate way to the forgiver's ceasing to hold the offense against the offender. The second is

that the change of judgment should be related appropriately to the kinds of reasons we think people can forgive for (when they do forgive for reasons). Like Murphy and others, I agree that one might forgive for a variety of reasons, including the recognition of past suffering, the recognition of current remorse and apology, and the recognition of a certain kind of moral luck. Sometimes the last idea is expressed as "there but for the grace of God go I." The fact that I might have done the same thing in the same circumstances, while not exonerating of the offender (just as it wouldn't be for me, were I to have done it), is a fairly common reason for not holding the offense against the offender.<sup>27</sup> Whether and when this is a *good* reason, or an overriding one, I will leave open for now. The crucial point for our purposes here is that not holding it against someone can be done for all sorts of reasons. Might it even be done for reasons such as the kind suggested by Dr. Phil and other self-help practitioners, namely, for your own mental health? My view is that it is possible in principle; one might choose not to hold an offense against someone, and thereby release her from her personal debt, in order to help oneself. But being very speculative here, I think that at best this would be psychologically difficult—it would seem just as likely that what one would succeed at with one's own health in mind would be "moving on" or a decision not to dwell on the offense or allow it to affect one's behavior in certain ways. Still, I think that if one genuinely succeeded at releasing the person, and ceased to hold the offense against her, then one would have forgiven. Finally, it seems possible that one's forgiving is motivated by certain desires without one's having adopted the fulfillment of those desires as justifying grounds for forgiving.<sup>28</sup>

Given this variety of reasons for forgiving, it may seem as if there is no single change of judgment that must take place in order to forgive. But there is a kind of minimal unifying judgment that is offender-directed when one forgives on this model, and it is this: it is a good thing, or something to be pursued, to cease to hold the offense against him. Insofar as the forgiving is intentional, one must have a positive attitude toward the end state; one must see it as a good thing that the offender not be in your debt for the culpable offense, despite his having committed it.<sup>29</sup> In many cases, this will take a more specific form, depending on the reasons for forgiving. Typically, it might include your thinking that there is something about the offender that makes it fitting or appropriate to release him or her. Post-forgiveness, there is of course a change, as well: the forgiver no longer sees the offender as in his or her debt, as owing anything.

What role if any does understanding of the offense, and of the offender's point of view play? There are at least two kinds of ways one might take the offender's point of view: in a replaying of the offense itself, and in imagining their experiences either before or afterwards, which might involve suffering or remorse. Taking another's point of view regarding the offense can certainly help generate the idea of moral luck, and the related idea that you, too, might have done the same thing in the same circumstances, especially if we extend the circumstances back in time to the offender's formative conditions. Although it is possible that seeing the events from the offender's point of view could make one all the more resentful and less inclined to forgive, there is some good evidence that it often has the opposite effect, and also that perspective-taking plays a role in many cases of self-attributed forgiveness.<sup>30</sup>

Interestingly, if we imaginatively identify with the offender in a mental replaying of the offense, we may be more likely to shift our attributions of the cause of the offense from something internal about the offender to something external about the situation. This tendency can be explained by the Actor-Observer Bias, according to which we are more likely as actors ourselves to attribute causes of behavior to our own situations, while we are more likely as observers of the behavior of others to offer internal attributions for the causes of their behavior.<sup>31</sup> If we take another's perspective on events, we imagine ourselves as actors, and thereby are more likely to make attributions as we would if we had acted ourselves. Now it is natural for people to infer from the existence of external or situational causes that the agents are not responsible for their actions, or not *as* responsible as they seemed when we were making more internal attributions. In fact, psychologists themselves regularly make this inference, sometimes implicitly, sometimes explicitly.<sup>32</sup> But there is no logical entailment here; it may be that one is responsible precisely for failing to resist the situational pressures that feature centrally in the causal explanation of one's actions, for example. Similarly, an internal disposition might be the best explanation of someone's action, and yet she is not responsible at all (because it is something she has no control over, for example). Still, the fact that this is such a natural inference can help explain cases that may fall between forgiveness and excuse—or, rather, that are overlapping of the two categories. In some cases in which people claim to forgive, they also appear to be lowering the degree of responsibility they attribute to offenders, if not fully withdrawing their attributions of responsibility.<sup>33</sup> Excuse, like forgiveness, includes a ceasing to hold something against someone; but in full excuse, unlike forgiveness, one rejects one's original attribution of responsibility. The

fact that perspective-taking can have two effects at the same time explains why partial excuse and forgiveness may often go together: one comes to see the offender as less responsible, but one also comes to identify with the offender and see one's differences as relating solely to one's circumstances over which one lacks control oneself.

Perspective-taking can also relate to events and feelings of the offender that either happen or are expressed before or after the offense. The offender's past suffering, for example, can be a reason to cease to hold an offense against him, as in cases like Robert Alton Harris' perhaps. By identifying with the suffering, one might feel that the burden of the obligation he has incurred is simply one too many. This is not to say that the obligation he has incurred as a result of the offense is an obligation to suffer *per se*, however. Perhaps the offender would be happy to fulfill the obligation. But it is a kind of burden and one might conclude that the offender has already had too many. Something similar might be said about the remorse and pain one perceives when one takes the point of view of the repentant offender. Thus, perspective-taking, and understanding from the point of view of the offender can play a number of different roles in giving reasons to forgive on the debt-release model, even if perspective-taking is not strictly necessary for forgiveness.<sup>34</sup>

At this point, one might ask what role the cessation of resentment plays in forgiveness, since it is not explicitly captured in the idea of debt release and continued judgment of responsibility. Not surprisingly, as has already been emerging, there is just as much controversy about the nature of resentment as about the nature of forgiveness. Some, like Murphy, take it to aim at the suffering of the offender. But others, like Hieronymi, take it to be a kind of mental protest. In their later work, Garrard and McNaughton seem to suggest that there are a variety of aspects to it, and prefer to use "indignation" to refer to a kind of protest, while reserving "ill will" for a hostile emotion that involves wishing ill, or wishing that things go badly for the person, rather than wishing someone well.<sup>35</sup> Without entering into a full adjudication about how we either do or should use the word, "resentment," it is worth distinguishing between the two concepts just canvassed as well as a third. There is sometimes a desire to have a certain kind of power over another, or more generally, a desire that the power relationship change in one's direction. (This might be manifested in part by thoughts like, "who are *you* to do that to me?" and "you think you are entitled to treat me that way?") It may be that this desire is yet a third part of something that sometimes goes under the heading of "resentment," and it is easy to see how this is related to the idea that one no longer holds another to the obligation incurred. It is true that one might

still have such a desire after having forgiven. Yet, at the same time, there is a tension in having seen the good in releasing someone from the obligation and having done so, and at the same time desiring a shift in a kind of power in one's own direction. Thus, though there may be residual feelings of this aspect of resentment after one has forgiven, they do not sit comfortably with it on the debt-release model. Thus, the fact that forgiveness is associated with a reduction of resentment in this aspect is well explained on the model.

What of the other two concepts mentioned earlier? Let us take them in turn. While I do not believe one need start with a desire for suffering to forgive, it may very well be a part of many instances of forgiveness. And clearly such a continued desire is at the least a sign that one has not forgiven. This can also be explained on the debt release model by reflecting on the typical reasons there are for forgiveness. If one is moved to forgive on grounds of a person's suffering (either because of events unrelated to the offense or because of their own remorse), then there is at the least a tension in wishing them continued suffering. Even if one is moved to forgive on grounds of understanding the event in ways that allow one to reflect on moral luck (in an abstract way or in the "there but for the grace of God go I" way), it seems that there will at the least be a tension between wishing suffering on others, while identifying with them. Perhaps it is hardest to explain a relationship between a diffusion of feelings of ill will and debt release in cases in which the forgiveness is motivated by considerations having to do with one's own mental health or for reasons unrelated to the offender's situation. And perhaps this very fact explains why Murphy and others have thought that one does not have true forgiveness if not done for the right sorts of reasons. But I think it is still possible to resist this reaction by showing that on the one hand, such cases might be very rare because psychologically difficult, and that where such cases exist, they are not the paradigms of the most *admirable* kinds of forgiveness. Feelings of ill will would seem to be just as destructive to one's mental health as holding offenses against others; thus if one is concerned primarily with one's mental health, it might be hard to succeed precisely in ridding oneself of the latter but not of the former. And even if one could manage it, we may simply have to say that though this is forgiveness, it is not of the most admirable sort.

Finally, we have the idea of protest.<sup>36</sup> Perhaps of all of the concepts that might go under the heading of resentment, this is the one whose elimination it is easiest to see as not required to forgive. Suppose one harbors no feelings of ill will, one has no desire for any shift in power relations, one has released the offender from any obligation with respect to

the wrong committed. One has forgiven. It seems conceivable that one nevertheless protests the offense and the message it carries.

In these different ways, the debt release model can accommodate the central idea that at least many attitudes that are sometimes subsumed under "resentment" are related to forgiveness in intimate ways. In the case of desires relating to changes in power, there is a kind of irrationality in forgiving while maintaining them (this may indeed be the case when one feels feelings of resentment at an old teacher, despite having forgiven, for example). In the case of feelings of ill will, the tension between maintaining them while having forgiven may be somewhat less straightforward, but is nevertheless present in many cases. And in the case of feelings that express protest, they may very well be consistent with having forgiven, but this is not a disadvantage for the view.

Finally, the last of the *desiderata* is an answer to the normative questions. Very briefly, I think the model is at least consistent with what I think are the most plausible answers. Sometimes, one simply ought to forgive: the offense was not so serious, the offender is remorseful and apologizes and offers to pay for the loss and more. Even here, though, I do not think this entails that the offender is entitled to demand that you do; he does not have a right to it. This is simply a case in which you have an obligation to do something for which no one else can claim a right. In other cases, perhaps there is good enough reason not to forgive. The offender is unrepentant, the offense was serious, the offender is likely to offend again against you and others.<sup>37</sup> There may be overriding reasons in such a case to hold the offender to the obligation. Finally, there may be cases in which forgiveness is not required, but is supererogatory. Such cases might include ones in which the offender has committed a serious offense and is unrepentant, but even at some cost to you, you find reason, perhaps in his past suffering, to release him from his obligation. The possibility of required, and, *a fortiori*, of supererogatory forgiveness explains why it is often claimed that forgiveness, or, more precisely, having a disposition to forgive, is a virtue to be cultivated. These answers obviously deserve more discussion than I can provide here. I hope here simply to have sketched the kind of approach the model makes possible.

## 5. Conclusion: Forgiveness and Freedom

The model I have developed here retains the idea, suggested in Strawson's paper, that forgiveness is rightly counted as a reactive attitude. When one forgives an offense, one is committed to the proposition that the offender acted freely and was responsible for the offense. If the skeptical

arguments turn out to be sound, then forgiveness, along with the rest of our practices of responsibility, will turn out to be unjustified. We could still understand and excuse, and in fact, we would be right to do so after making any initial attribution of responsibility, but forgiving would be seen to rest on an illusion. Thus, forgiveness remains one of the central aspects of human life at stake in the debate.

This picture contrasts with a model of forgiveness that omits any requirement of an attribution of responsibility. Derk Pereboom (2009) puts forward this sort of account, according to which forgiving is a kind of decision to continue the relationship, despite one's having been wronged and recognizing that one's relationship has been impaired as a result (pp. 183-84). Making this kind of decision to restore or continue the relationship can amount to the retraction of a change in behavior premised on the impairment of a relationship, such as shunning or avoiding. Yet Pereboom does not argue for this account; rather, he suggests that even if such an account is not really an account of *forgiveness*, we can still retain valuable features of it. For example, if we are not responsible agents in a robust desert-entailing sense, then we can still make the choice to restore previously impaired relationships in the face of genuine apology and moral resolve. While reconciliation is often associated with forgiveness, it is neither a necessary nor sufficient condition for it on my account. Thus, the roles played by each of forgiveness and this special kind of relationship restoration will be significantly different. Whatever we ultimately conclude about the details of their functional roles in human life and about their relative value, the debt release model developed here offers further support for the idea that an attribution of a robust sort of responsibility is a central part of forgiveness and so of a central part of human life, at least as currently lived.

At the same time, the debt release model also brings out a variety of dimensions of blame, ones important for the debate itself. It shows how some considerations related to forgiving, like the recognition of moral luck, can give us reasons not to blame in ways *related to holding to personal obligation*. At the same time, there may remain reasons to blame in other robust ways. These can include judgments of desert-entailing blameworthiness, the making of moral demands and protests, and perhaps even feelings of certain kinds of resentment. They might also include an approval of institutional punishment. While incompatibilists often take certain implications of determinism to undermine blameworthiness, the consideration of forgiveness here illustrates how at least some such implications, including moral luck in one's circumstances, might in fact be taken to provide reasons not to blame only in certain ways. If there is to

remain a genuine distinction between excusing and forgiving, and if forgiveness is at least appropriate in some circumstances, then considerations in favor of forgiving cannot undermine all aspects of appropriate blame. These are big "ifs," and I have not here assessed any of the arguments for general skepticism about freedom and responsibility. My aim here has been instead to point to some interesting parallels between forgiveness and the debate about freedom and responsibility, and to sketch the framework of an account of forgiveness that I hope can shed some light on the nature of both.<sup>38</sup>

## Notes

<sup>1</sup> See Larsen et al. 2012.

<sup>2</sup> See, for example, Kolnai 1973 and Calhoun 1992.

<sup>3</sup> For example, see Pereboom 2001 and 2009.

<sup>4</sup> A related, but distinct, way in which the nature of forgiveness is relevant to the debate about human freedom and responsibility is in arguments over whether one can hold responsible for an instance of wrongdoing without judging that sanctions are a warranted response to it. I discuss this issue in Nelkin 2011.

<sup>5</sup> The word "forgiveness" is used in a number of different ways (e.g., in "debt forgiveness"), a point to which I return in section IV. For now, I mean to locate forgiveness as in the family of the reactive attitudes, but recognize that even here, the word, 'forgive' may be used more or less loosely and yet still quite aptly. At the same time, recognizing this important point is consistent with its being useful to explore the contours of a particular phenomenon that plays an important role in human life, and that is connected in intuitive ways to other familiar concepts such as resentment and apology.

<sup>6</sup> Because of limitations of space, I here focus primarily on three recent accounts that allow for the drawing of certain key insights especially relevant to the view to be developed in section IV. But I note some important ways in which the account is influenced by, or contrasts with, some of a number of other rich accounts along the way.

<sup>7</sup> Butler 1846. See also Newberry 2003 for an interpretation of Butler that nicely highlights this point.

<sup>8</sup> At least not in the way that is a moral virtue, he adds. Murphy and Hampton 1988, p. 24.

<sup>9</sup> See Watson 1987 for a detailed and illuminating discussion of this case.

<sup>10</sup> Murphy and Hampton 1988, pp. 39-40.

<sup>11</sup> Hieronymi here responds to the kind of objection Calhoun (1992) raises for views that take forgiveness to be obligatory.

<sup>12</sup> Watson 1987/2007.

<sup>13</sup> See 2003 and 2010, p. 90.

<sup>14</sup> See Fischer and Tognazzini 2011 for a very helpful taxonomy of a whole series of what they call "stages" of responsibility, illuminating different ways in which

we might appropriately blame and showing that the conditions for each might be different. See also Scanlon 2008.

<sup>15</sup> If God can forgive, and yet is incapable of having feelings of ill will, then God, too, would be a counterexample. There is much debate about whether the notion of forgiveness by God and forgiveness by humans is the same, and I will set that aside here.

<sup>16</sup> One might disagree on this point by claiming that there could be contexts in which we believe it is unparadoxical to think to oneself, "I forgive him, but I think he should apologize." As I argue in section IV, once we distinguish between beliefs about what others ought to do and what we demand from them, we can account for such coherent thoughts and still agree that demanding an apology is a sure sign of not having forgiven.

<sup>17</sup> See Hieronymi 2001 for a similar point.

<sup>18</sup> I defend this kind of account in Nelkin 2011, but here develop it in more detail.

<sup>19</sup> See Murray and Rea 2011 and the references therein to Anselm 1998.

<sup>20</sup> It is worth noting that Garrard and McNaughton cite Swinburne (1989) who is himself offering an explicitly Christian model.

<sup>21</sup> What is it to readjust your relationship? Twambley writes that while your esteem for the person might have gone down in light of the offense, "you do not hold the offense against him, you 'bear him no grudge,' his action is no longer such an impediment to your relationship...you *re-accept* him" (1976, p. 89). This goes beyond what seems to be strictly required in a debt release model in requiring that the offense not serve as an impediment to the relationship. But as many have noted (see, for example, Murphy and Hampton 1988), one might forgive without reconciling. The example most often cited is that of an abused wife who forgives her husband, but nevertheless pursues a divorce. Thus, it is important to separate the debt-release model from a requirement of reconciliation or relationship repair.

<sup>22</sup> There is a further question as to whether the harm caused—which may be a matter of consequent moral luck, and not in the control of the offender—is part of the basis of one's obligation to make up for the wrong (as opposed to the obligation to pay for the loss). But I will set that aside here.

<sup>23</sup> Interestingly, Twambley, who advocates this kind of model, criticizes competing views on the grounds that they make forgiveness something that is "earned" and "obligatory". But he seems to be unaware that the framework of the debt-release model alone seems subject to the same kind of challenge. On the other hand, Swinburne (1989) acknowledges the challenge. His solution is to suggest that the victim has the power to maintain the "guilt" of the offender within limits (and "for a while"), but that there are some cases in which the offender has apologized, repented, made reparations, and given penance, in which the guilt can "vanish" despite his not having been forgiven (1989, p. 88).

<sup>24</sup> It is also possible, on this view, that the victim sets the obligation lower than she should. In this case, it might still be that the offender ought to do more than this personal obligation requires—not because she owes it to the victim in virtue of this obligation, but simply because she ought to do it.

<sup>25</sup> Although once we get clearer about the nature of resentment, this may become an option once again. (See the discussion of different aspects of resentment below.)

<sup>26</sup> See also Griswold 2007 and Allais 2008, for other well-developed accounts that take a change in judgment to be essential. Like the account developed in the text, Allais' account is sometimes described in terms of ceasing to hold the offense against one. But Allais explicitly distinguishes her interpretation of this from a kind of debt-release model, and instead understands forgiveness, and ceasing to hold it against someone in terms of seeing the agent differently. "On my account, holding something against someone involves lowering the way you affectively esteem or regard her as a result of her action: your attitude toward her is more negative than it would be if you did not see the action as counting in the evaluation of her as a person" (20098, p. 56). Thus, forgiving, on this view is a kind of "wiping the slate clean." While, I agree that some change of judgment is entailed by forgiving, I do not think this particular change of judgment is necessary for (and, *a fortiori*, constitutive of) forgiveness. It seems to me that one might continue to regard one's offender more negatively as a result of the offense (say, in the case of an act of abuse on the part of a spouse) and, as a result, leave the relationship because of that very act, but one might still forgive. I believe that the debt release model accounts more naturally for this possibility.

<sup>27</sup> See Takaka 2001.

<sup>28</sup> This is not to find all of Dr. Phil's words on this topic coherent. Forgiveness might be about you in the sense that your own good motivates you; but it can't be "all about you" (emphasis mine) because you have to take attitudes that have the offender as your object, and in a way to be described, change your attitude toward that person. Oscar Wilde's prescription, funny as it is, might also be possible in principle, but even harder psychologically. For if you conceive of the offenders as enemies, it might be very hard for you at the same time to see releasing them from their obligation as a good thing.

<sup>29</sup> This is consistent with the offender owing you compensation for the loss or harm incurred.

<sup>30</sup> For some suggestive results concerning a link between perspective-taking and forgiveness, see Darby et al. (in preparation) and Takaka 2001. The fact that taking the offender's point of view might make one angrier is offered by Hieronymi (2001) as a reason against making it a necessary condition of forgiveness, or as the centerpiece of a sufficient condition as Novitz (1998) does. Griswold (2007) takes a certain sort of "entering sympathetically into the situation and self of the offender" to be a necessary, but not sufficient condition of forgiving, for the very reason that taking the offender's point of view might increase the victim's resentment (p. 57).

<sup>31</sup> For the classic discussion of the Actor-Observer bias, see Jones, E.E., and Nisbett, R.E. 1971.

<sup>32</sup> See, for example, Shepperd, J., Malone, W. and Sweeny, K. 2008.

<sup>33</sup> See Darby et al. (in preparation).

<sup>34</sup> Interestingly, in offering a specifically Christian model of forgiveness Marilyn Adams (1991) argues that forgiveness essentially depends on a different kind of perspective-taking: namely, that taking the perspective, to the extent possible, of God. Seeing the situation from God's perspective allows one to see the offender in a different light. On the debt-release model, which is compatible with a variety of reasons for forgiving, this could be one among the possible reasons for forgiveness.

<sup>35</sup> See Garrard and McNaughton 2010, p. 23 and p. 91.

<sup>36</sup> Outrage might be interestingly different, and though I set it aside here, note that it deserves its own discussion.

<sup>37</sup> Because Garrard and McNaughton see forgiveness primarily in terms of overcoming ill will, they conclude that there is always good reason to forgive if one does so for the right reasons. But once we shift to a debt release model, and see the overcoming of other aspects of resentment as the more immediate consequences, we do not have the same reasons for a universal endorsement of forgiveness. Similarly, but for somewhat different reasons from Garrard and McNaughton, Holmgren (2012) argues that unconditional forgiveness is always called for (consistent with the forgiver's having addressed the moral wrong). But here it may be that the disagreement is not straightforward, since on her view, forgiveness is always "appropriate and desirable from a moral point of view" (p. 10), and this is not the same as something one is morally obligated to do.

<sup>38</sup> A version of this paper was presented at the Bowdoin Workshop on Free Will in October 2011, and I am very grateful for the excellent commentaries by Laura Ekstrom and Daniel Speak, and thank all of the participants for their very valuable feedback, including Mark Balauer, Randolph Clark, Peter Graham, Eddy Nahmias, Scott Schon, Neal Tognazzini, and Manuel Vargas. I also benefited greatly from input from my colleagues in the UCSD Chancellor's Interdisciplinary Collaboratory on the Nature of Forgiveness and Its Role in Cardiovascular Health (2010-2012)—Nicholas Christenfeld, Ryan Darby, Christine Harris, Britta Larsen, and Per Milam. Finally, I am indebted to Eve Garrard, Margaret Holmgren, David McNaughton, Derk Pereboom and Sam Rickless for their very helpful and insightful comments on previous drafts.

## CHAPTER NINE

### MORAL RESPONSIBILITY, FORGIVENESS AND CONVERSATION

BRANDON WARMKE  
AND MICHAEL MCKENNA

#### Introduction

All but the hardest of us have forgiven someone for something. Many of us, probably most, have not forgiven someone we should have, and have forgiven someone we shouldn't have. Based on this very simple observation about our everyday lives, we note two questions one might press, one analytical, another normative: First, what is forgiveness? Second, what norms govern the propriety of it?

In this essay we shall attend primarily to the first of these questions, but our hope is that what we say about it will also be suggestive of how one might take up the second. As for the first, we shall answer it within the context of a theory of moral responsibility. Naturally, for this reason, we shall restrict our attention to what might awkwardly be called *moral forgiveness*. Just as there might well be other kinds of responsibility other than moral, such as legal, practical, aesthetic, or even etiquettical, so too on our account there might be other kinds of forgiveness.<sup>1</sup> Those who disagree on this score—that is, those who think all forgiveness is moral—can handily set aside their dispute with us on this point, since, if all forgiveness is moral, then we can be taken to be offering an unrestricted rather than a restricted treatment, in which case, we'll happily drop the modifier 'moral' as it would be an obvious redundancy. We'll begin by sketching a conversational theory of moral responsibility. Then, drawing upon this theory, we shall develop an account of forgiveness, one that, admittedly, will not be exhaustive (nothing short of a book-length treatment could be). We'll close by discussing benefits of this approach and suggesting how one might begin to settle the normative question.