

Articulate forgiveness and normative constraints

Brandon Warmke

Department of Philosophy, University of Notre Dame, Notre Dame, IN, USA

ABSTRACT

Philosophers writing on forgiveness typically defend the Resentment Theory of Forgiveness, the view that forgiveness is (or crucially implicates) the overcoming of resentment. Rarely is much more said about the nature of resentment or how it is overcome when one forgives. Pamela Hieronymi, however, has advanced detailed accounts both of the nature of resentment and how one overcomes resentment when one forgives. In this paper, I argue that Hieronymi's account of the nature of forgiveness is committed to two implausible claims about the norms bearing on forgiveness. Her account is highly instructive, however, for it brings into relief how certain intuitive views about the norms of forgiveness should be used to constrain our theories about its nature. I conclude by defending this methodological proposal.

ARTICLE HISTORY Received 27 May 2015; Accepted 25 September 2015

KEYWORDS Forgiveness; resentment; blame; apology; obligation; control

1. The resentment theory of forgiveness

Philosophers writing on forgiveness have tended to converge on what we may call the Resentment Theory of Forgiveness:

(RTF) Forgiveness is (or at least crucially implicates) the overcoming of resentment.

According to this theory, to forgive someone for her having wronged you is (roughly speaking) to overcome the resentment you have towards her because of what she did to you.¹ Stephen Darwall, for example, claims that '[t]o forgive is, roughly, to forbear or withdraw resentment' (2006, 72). Darwall is not alone in ascribing to the RTF, which has so dominated recent discussions of forgiveness that it has been described as the 'standard view'² and 'received orthodoxy'.³

As a theory of the nature of forgiveness, the RTF encounters two problems. The first is that there is little agreement as to how resentment should be understood.

CONTACT Brandon Warmke  bwarmke@nd.edu

© 2015 Canadian Journal of Philosophy

Some think of resentment as a 'hostile feeling' which aims at inflicting harm on the wrongdoer (Garrard and McNaughton 2002), while others hold that it is a kind of 'moral protest' (Hieronymi 2001), while still others claim that resentment is the paradigmatic sort of 'moral anger' (Hughes 1993). Furthermore, whereas some describe resentment as a 'vindictive passion' (Murphy 2003, 16), others claim that resentment 'need not entail motivation to retaliate for the wrong' (Holmgren 2012, 31). It is difficult to know what exactly these accounts of resentment amount to, what kinds of relations hold between them, and, perhaps most importantly, which account is to be preferred.

A second problem facing the RTF is that while there is widespread agreement that forgiveness crucially implicates resentment (whatever it is), there is no such agreement about what must *be done* with resentment in order to forgive: What does it take to overcome resentment in a way that qualifies as forgiveness? Some have claimed that to forgive is to 'abandon' or 'forbear or withdraw' resentment.⁴ Still others claim that forgiveness involves 'letting go' of or the 'elimination of' resentment.⁵ P. F. Strawson claimed that, in forgiving, one 'forswear[s] resentment' (1962, 76). Rarely is much more said about the machinery of these proposed changes to one's emotional life.

In her admirably deep paper, 'Articulating an Uncompromising Forgiveness' (2001), Pamela Hieronymi recognized and sought to address these two problems by providing both a theory of the nature of resentment and a detailed account of how one overcomes resentment when one forgives. While recent discussions of Hieronymi's account of forgiveness have contested her view on the grounds that it misconstrues the *nature* of forgiveness (e.g. Zaragoza 2012 and Nelkin 2013), I will argue that her theory of forgiveness results in untoward consequences as regards the *norms* of forgiveness.⁶ In Section 2, I set out the main contours of Hieronymi's proposal. In Sections 3 and 4, I argue that her account commits us to two implausible theses regarding the norms of forgiveness. In Section 5, I reflect on these problems and suggest that they are instructive for theorizing the nature and norms of forgiveness.

2. Hieronymi on forgiveness

One of the primary worries animating Hieronymi's work on forgiveness is a frustration with the state of the philosophical literature, which is dominated by 'the all-too-common habit of talking about resentment and anger as things to be manipulated – to be "banished," "destabilized," "dissipated," "dispelled," "vanquished," "conquered," or "removed" – rather than as attitudes sensitive to one's judgments, subject to rational *revision*' (2001, 535).⁷ In Hieronymi's terminology, such views of forgiveness are *non-articulate*. As one of Hieronymi's commentators, Kevin Zaragoza, has put it, non-articulate versions of the RTF understand the overcoming of resentment as the elimination of:

resentment by indirect or nonrational manipulations, such as imaginatively identifying with the wrongdoer's point of view, thinking about the positive aspects of one's past history with the wrongdoer, etc. Such non-articulate accounts basically involve taking steps in order to accomplish a desired effect of eliminating resentment. (2012, 609)

Hieronymi rejects non-articulate accounts of forgiveness. Why? As Zaragoza has thoughtfully put it, one alleged problem with non-articulate accounts is that they face a dilemma. According to non-articulate accounts, forgiveness must be identified with either: (1) the non-rational manipulations or processes by which one eliminates resentment; or (2) the loss of the resentment itself (or perhaps the conjunction of the processes leading up to the loss of resentment and the loss of the resentment itself) (2012, 609). However, the first horn of the dilemma requires the rejection of the 'widely accepted and highly intuitive contention that forgiving requires the elimination of resentment' (Ibid.). One could, after all, engage in the right kinds of preparatory processes and manipulations without *actually* eliminating the resentment. One might hope and try to rid oneself of resentment, but there are no guarantees. So taking this horn would be to claim that one could forgive without actually eliminating resentment, which, for a defender of the RTF, is obviously an undesirable result.

Alternatively, the second horn of the dilemma has the result of unacceptably diminishing (or eliminating altogether) the kind of control that is commonly thought to be essential to forgiving. Forgiving is something an agent *does* – it is not something that *just happens* to someone. Yet according to non-articulate accounts, forgiveness is *not* something one does; it is rather the 'hoped-for result of other steps that we take' (Zaragoza 2012, 610). In other words, if non-articulate accounts are correct, we don't make it the case that we forgive by *acting*. Instead, we engage in other activities that may or may not accomplish our goal of bringing about the desired result: the elimination of resentment. To equate forgiveness with the elimination of resentment on a non-articulate account is therefore to put forgiveness too far outside the reach of our agency. This would be true even if we thought of forgiveness as the *conjunction* of the two horns: we would still lack the right kind of control over our elimination of resentment. An adequate account of forgiveness, therefore, must avoid being non-articulate.

As Hieronimi sees things, our best theories of forgiveness should therefore be *articulate*. 'An articulate model of forgiveness,' she writes, 'would explain what revision in judgment or change in view would serve to *rationally* undermine justified resentment in something like the way an excuse undermines indignation' (2001, 535).⁸ The general point here is that an account of forgiveness should make use of the resources provided by the fact that emotions like resentment (or indignation, fear, or jealousy) are subject to rational revision, which can thereby undermine them and bring about (in properly functioning psyches) their elimination. Take, for example, the case of indignation. You might see what you take to be a thief removing a pear from a neighbor's fruit stand. Your collection of

judgments about this situation gives rise to your feeling indignation. But were you to discover that the putative thief was not a native English speaker and misread a sign, thinking it said 'FREE PEARS,' you might revise some of your judgments. (Perhaps you revise your judgment that this person acted out of a bad quality of will or lack of due regard for the owner of the fruit stand.) Such a revision would (typically) bring about an elimination of your indignation.

A similar kind of story can be told about forgiveness: our theories of forgiveness should be able to explain how it is that a revision of a judgment might bring about the elimination of resentment. So just like our theories of excuse and justification could explain how revisions to our judgments can bring about an elimination of indignation, so too our theory of forgiveness should be able to explain how a revision in judgment can also bring about an elimination of resentment. To provide such a theory would be to provide an *articulate* account of how resentment is overcome when one forgives.⁹

But not just *any* articulate account of how forgiveness crucially implicates the elimination of resentment will be acceptable. Our best theories of forgiveness must also be, according to Hieronymi, *uncompromising*. An uncompromising account of forgiveness is one that is able to explain how forgiveness could be a way of eliminating resentment that is distinct from other ways of eliminating resentment: justification, excuse, and condonation.¹⁰

Take the case of justification. Suppose you see what you believe to be a stranger stealing a pear from your fruit stand. Resentment wells up within and you steel yourself to go confront the alleged thief about her despicable deed. But then you realize that you are mistaken: the stranger wasn't taking a pear from your fruit stand, but rather took the pear from an adjacent stand owned by the stranger herself. Upon realizing that the stranger did nothing wrong and was therefore justified in taking the pear, your resentment fades away. But even though you have eliminated your resentment by revising a judgment, this does not mean that you have forgiven the stranger. It would be an undesirable consequence of an articulate account of forgiveness if it allowed that *any* elimination of resentment that was the result of a revised judgment counted as forgiveness.

The lesson here for defenders of the RTF is that whatever story is to be told about the elimination of resentment, that story must be different than the one we tell about what happens when one justifies, excuses, or condones. The name of the game is to figure out which of the judgments that undergirds resentment is the unique judgment that is revised when one forgives. To give an articulate *and* uncompromising account, then, is to provide an account of two things: (1) the judgments that undergird resentment and (2) the judgment the revision of which is unique to forgiveness, such that when one revises *that* judgment (and thereby eliminates resentment), one forgives. Hieronymi provides just such an account, to which we now turn.

Hieronymi claims that in order to resent, minimally one must both judge and possess some sort of commitment to each of the following:¹¹

- WRONG: The act in question was a serious moral offense.¹²
- RESPONSIBLE: The wrongdoer is morally responsible for the act in question, and as such may legitimately be blamed.
- SELF-WORTH: You, as the one wronged, ought not be wronged.

Although these three judgments undergird resentment on Hieronymi's view, she argues that none of them is the unique judgment that is revised when one forgives. This is because to revise WRONG would be to *justify* the author of the action since one would no longer judge the agent to have done wrong. To revise RESPONSIBLE would be to *excuse* or *exculpate* the agent since one would no longer judge the agent to be blameworthy or morally responsible. And finally, to revise SELF-WORTH would be to condone what the agent did since one would no longer judge that one should not be the victim of a serious moral offense. All of these revisions can be ways of eliminating resentment, but because a desirable account of forgiveness must be uncompromising, we must identify some *further* judgment that undergirds resentment such that when *it* is revised, we may have a case of forgiveness.

This key fourth judgment can be brought into relief by attending more carefully to the nature of resentment. Hieronymi understands resentment as a form of protest. In resentment, she writes, 'the victim protests the trespass, affirming both its wrongfulness and the moral significance of both herself and the offender' (2001, 530). In her view, victims have something to protest because past wrongs *persist as present threats* to their victims (546).¹³ Past wrongs claim 'that you can be treated in this way, and that such treatment is acceptable' (Ibid.). When you knowingly and willfully deceive me, for example, this shows that you think of me as the kind of person for whom such deception is an acceptable mode of treatment. Consequently, the claim made by your past action stands as persisting threat – a continuing statement that I can be treated that way.

To resent, then, is to protest what a past action presently signifies.¹⁴ Resentment protests these claims made by the wrongdoer; it affirms the wrongness of the action and one's self-worth against the threat posed by the injury. For Hieronymi, then, these foundational three judgments 'together imply (other things equal)' a fourth judgment that she says, 'grounds resentment' (548):

- THREAT: The event in question makes a threatening claim.

This is the key judgment undergirding resentment, and it is the judgment that explains what happens when one forgives: to forgive is to overcome resentment by revising this fourth judgment *while at the same time* retaining the first three. After all, these first three judgments must be retained if forgiveness is to be *uncompromising* (i.e. if it is not to collapse into justification, excuse, or condonation). To forgive is to judge that the past action no longer makes a threatening claim, which is to make a revision to THREAT. Such a revision to THREAT thereby *rationaly undermines* our resentment. And when we 'come to

see' that our resentment is unfounded, it 'will (in the well-functioning psyche) disappear' (535).

But how is THREAT rationally undermined? Here, Hieronymi calls our attention to how apologies often function. Past wrongs can stand as present threats, and one thing that apologies can do is to retract the threatening claim made by that past wrong. She writes: 'Once the offender *himself* renounces the deed, it may no longer stand as a threat [...] It has been cut off from the source of its continued meaning. The author has retracted his statement, and anger loses its point' (548). Apologies can, as it were, diffuse the threat in THREAT, thus rationally undermining it. Indeed, Hieronymi regards this as a consideration in favor of her account: it can explain both the function and normative force of apologies: apologies function to retract threatening claims, thereby altering what is rationally permissible for a victim to believe about the wrongdoer's past action. As Hieronymi puts it, apologies change 'the significance of the event,' and so therefore 'resentment loses its footing' (549).

To summarize: Hieronymi argues that any account of how forgiveness crucially implicates the overcoming of resentment ought to be *articulate* and *uncompromising*: it must identify a unique judgment that is rationally revised when one forgives. That key judgment is THREAT, which can be rationally undermined when wrongdoers apologize. Revising THREAT typically results in the disappearance of resentment. This is how forgiveness crucially implicates the overcoming of resentment.

3. Forgiveness and control

If you recall, one alleged problem with *non-articulate* accounts of forgiveness is that according to such views, forgiving must either be identified with the process that only sometimes leads to the elimination of resentment, or instead with the elimination of resentment itself (or the conjunction of the two). The former option retreats from the claim that forgiving is the overcoming of resentment (because all of one's efforts may not actually eliminate the resentment), and the latter options put forgiveness too far outside the victim's control (because the elimination of resentment is not something that an agent can control). One of the motivations for an articulate account, therefore, was to prevent forgiveness from turning into something that was accomplished by non-rational manipulations, the machinations of which rest largely outside of our control. But because revising a judgment is something that *we do* (instead of something that just *happens* to us), the thought was that an articulate account gives a forgiver sufficient control over her forgiveness.

The problem, however, is that a similar worry plagues Hieronymi's articulate proposal. To see how, notice that on her account, standard cases of forgiveness involve the following three features or stages:

- (1) The rational revision of some judgment that makes continued resentment rationally unfounded;
- (2) the 'com[ing] to see that our resentment is unfounded'; and
- (3) the consequent 'disappear[ance]' of resentment in the well-functioning psyche.¹⁵

The matter I now wish to consider is which, if any, of these aspects of forgiveness is under our voluntary control. Stage (1) describes a revision to a judgment we make about someone's past action: that it no longer constitutes a present threat. But it is not clear that this is something that is under an agent's voluntary control. We may revise a judgment without, for example, *deciding* to do so. Indeed, this seems to be how we revise many judgments, like the revisions we are constantly making about our physical surroundings. Yet we are not making these judgments by voluntary effort – about such matters the doxastic involuntarist is correct. When the previously threatening bear near your campsite scurries away and over the hill, you may revise your judgment that the bear poses a threat – but this need not be something you voluntarily do. The same might be said about receiving an apology: when given a sincere apology, you may revise THREAT without voluntarily doing so.

The same goes for stage (2). Our 'coming to see' that our resentment is unfounded need not be the kind of thing that we do voluntarily. It just may seem obvious to us. Just as our coming to see that there is no reason to continue to judge that the bear is a threat is not something we voluntarily do, so too, our coming to see that there is no reason to continue to judge that a past wrong action stands as a present threat is not something that we do voluntarily.

Stage (3) does not fit the bill either. The 'disappearance' of resentment in a well-functioning psyche need not be anything that we do by voluntary effort; indeed, the locution itself suggests that it is not. The fact your fear of the bear dissipates once you see that there is no reason to take the bear to be a threat is not under your voluntary control. In like manner, the fact that your resentment dissipates once you see there is no reason to take the past action as a threat is not something that is under your voluntary control. Furthermore, it is plausible that stage (3) is the *least* likely aspect of this process to involve voluntary control. And yet this aspect of forgiveness is the crux of any theory that understands forgiveness as the *elimination of resentment*.

Perhaps the rational revision at stage (1) is *more* within our control than is the resentment-elimination at stage (3), but then the heart of forgiveness would turn out to be a matter of belief-revision, not of the overcoming of resentment. And even if this is correct, it is not clear that we have more control over this belief-revision than we do over the kinds of indirect attempts at resentment-elimination that motivated Hieronymi to reject a non-articulate account in the first place. If one of the alleged problems of non-articulate accounts is that they present a picture of forgiveness whereby *forgiving* is only indirectly

achieved and therefore fails to capture the sense in which forgiving is something we *do*, then the Hieronymi strategy does not fare much better. On her account, forgivers look like passive machines that revise judgments in response to one's environment, thereby effecting emotional changes. One can forgive on such a view without voluntarily doing much of anything at all.

How did this happen? Hieronymi set out by rejecting the view that resentment is something to be 'manipulated,' 'dissipated,' or 'dispelled.' But this seems to be the kind of view at which we ultimately arrive – a view according to which non-voluntarily acquired revisions to our judgments can result in a 'disappearance' of resentment. It is of course true that in her view whether we resent or not is typically sensitive to the judgments that undergird resentment. But this fact alone does not put forgiveness under our voluntary control. The correct diagnosis, I think, is that Hieronymi's discussion was originally motivated by attending to mechanisms by which resentment is eliminated in case of justification and excuse. If you recall, Hieronymi's account gains traction from the thought that in cases of justification and excuse, resentment can be rationally undermined and eliminated by the revision of certain judgments (e.g. that one did wrong; that one is blameworthy). This lesson was then applied to forgiveness. Ostensibly, the only difference between the three cases is the specific judgment that is revised. Like justification and excuse, it was claimed, forgiveness can be accomplished by the rational revision of a judgment that undergirds resentment.

The problem, however, is that to 'justify' or 'excuse' is, as far as I can tell, typically not something that one does voluntarily.¹⁶ When we justify or excuse, we come to recognize certain facts, such as the fact that an agent has not done wrong after all, or that if she has, she is not blameworthy for having done so. In justifying and excusing, we come to revise certain judgments, which tend to have diffusive effects on attitudes like resentment and indignation. Hieronymi is correct here. But if we think of forgiveness along these same lines, then it is no surprise that we get a theory that depicts *forgiving* as something largely outside of our voluntary control. This is what we should expect. The mistake, then, is to think that forgiveness operates like justification and excuse. Forgiving *is* something we do voluntarily.

If this is correct, then it may turn out that *both* articulate and non-articulate accounts of the RTF fail to give forgivers voluntary control over their forgiveness. In that case, we should seriously rethink the assumption guiding both of these views in the first place: that forgiveness must crucially implicate the overcoming of resentment. If paradigmatic cases of forgiveness typically require some sort of voluntary control and yet it is also true that we typically lack such control over our emotional lives, then this would be a good reason to be skeptical about views of forgiveness that crucially implicate certain sorts of emotional changes.

Perhaps, however, this has all been too quick. For while I have argued that Hieronymi's account does not secure forgivers voluntary control over their forgiveness, I have not said anything about why such voluntary control is desirable

in the first place. Maybe we can get everything worth wanting out of a theory of forgiveness without positing anything more substantial than the kind of rational control that Hieronymi has in mind – the ability to revise judgments about wrongdoers in a way that typically results in the disappearance of resentment.

I think that this view is mistaken, however, and I will now argue that our theories of forgiveness will be impoverished unless forgiveness is (at least in the paradigmatic cases) something that is under our voluntary control. The line of argumentation I shall pursue here involves showing that much of our common sense thinking about the *norms* of forgiveness requires that forgiveness be the kind of thing that is largely under our voluntary control. So consider the following argument:

- (1) Some of the norms bearing on forgiveness are ought-norms, norms that concern whether a victim ought to forgive (or ought not to forgive) the wrongdoer.

In at least some cases, one ought (or ought not) to forgive. This is not to say, of course, that forgiving is always a matter of doing what one ought or ought not to do. Sometimes forgiving may be a matter of permission, supererogation, or appropriateness or fittingness. But as Jeffrie Murphy writes, ‘sometimes it is not merely permissible that I forgive but that I *ought* to forgive and can be properly criticized if I do not’ (1988, 29). This thought is also commonly expressed in the ways we talk about forgiveness. For example, we say things like:

You really should forgive your wife for forgetting your anniversary, Dan. How long has it been, three months?

I hope you aren’t thinking of forgiving your boss so soon, Tina! You really shouldn’t – sexual harassment is a serious wrong!

Son, here’s some fatherly advice: never forgive someone unless they apologize first.

Now consider:

- (2) The ought-norms bearing on forgiveness sometimes presuppose that compliance with those norms is typically within the victim’s voluntary control.

In support of premise (2), notice that when ought-norms are invoked as they are in the above examples, it would be natural for those on the receiving end of such exhortations to think that what they are being asked or encouraged to do is something that they could either choose or refuse to do.¹⁷ Indeed, philosophers commonly describe forgiveness as an ‘elective response’ to wrongdoing (e.g. Calhoun 1992, 81), noting that sometimes this can instead involve a ‘refusal’ to forgive (Richards 1988, 95).

- (3) Therefore, some of the norms bearing on forgiveness sometimes presuppose that whether one ought to forgive or not forgive is within the victim’s voluntary control.¹⁸

The upshot of this argument is that, if sound, it shows that if we reject the view that forgiveness is at least sometimes under the victim's voluntary control, we must also reject the view that ought-norms sometimes bear on forgiveness. Adopting Hieronymi's account of forgiveness, therefore, comes at a theoretical cost: it would force us to revise at least some of the common sense ways we think of the norms bearing on forgiveness.

One objection to this argument is simply to reject the claim that ought-norms bear on forgiveness at all. If there are no ought-norms that bear on forgiveness, then we are not sacrificing anything by adopting a view of forgiveness that cannot account for them. But I cannot see any pre-theoretical reason for rejecting forgiveness ought-norms. Indeed, for those philosophers who explicitly reject such ought-norms, the reason seems to be simply that their account of forgiveness cannot account for them. Charles Griswold, for example, writes that 'because the sentiments are not wholly at the command of the will, the forswearing of resentment cannot be obligatory let alone subject to coercion' (2007, 68). But whether we should account for forgiveness in terms of emotional change is exactly what is at issue, or so I allege. And furthermore, there are good pre-theoretical reasons in favor of embracing these ought-norms, most notably because rejecting them would require a costly revision both to our practice and our thinking about forgiveness.

A second objection to this argument grants that it is sound, but claims that, even so, this does not cut ice against Hieronymi's view of forgiveness. Here, the critic can argue that the ought-norms bear on some course of action that the putative forgiver *can* engage in by an exercise of voluntary agency, and that this exercise of agency then initiates some psychological processes that typically result in the elimination of resentment. One might, for example, decide to go to counseling, or decide to give up and disavow resentment. Such voluntary behaviors may then lead to the revision of THREAT and the elimination of resentment. Therefore, so the objection goes, we can retain our commitment to the RTF and yet still secure both forgiveness ought-norms and voluntary control over some crucial aspect of the forgiveness process.

The problem with this objection is that it forces us either to relocate the ought-norms (so that they do not bear on forgiving itself) or instead to reject the view that forgiving is the overcoming of resentment. For if the ought-norms are to bear on something that is under our voluntary control, they can no longer bear on one's *forgiving* (if, that is, forgiveness is the overcoming of resentment). Instead, they bear on something like one's *trying* to forgive. On such a view, what one ought to do forgiveness-wise (and the thing one can do via an exercise of agency) is to try to forgive. But as Zaragoza himself noted in discussing the non-articulate accounts, this seems to get the norms of forgiveness wrong:

[On non-articulate accounts, f]orgiveness would not then be the act of the agent but rather a hoped-for result of other steps that we take. This divorce of forgiveness from our agency and control is certainly counterintuitive. It is also ethically

troubling. We standardly think of forgiving as a virtuous or good act. People are at least sometimes praised for forgiving, and may be blamed for being unforgiving. (2012, 611)

We sometimes blame and praise others for forgiving or not forgiving, and we issue normative assessments based on whether they (or we) forgive – not simply whether one has *tried* to forgive. Now of course, there may be ought-norms that bear on the trying too, but the question is whether those are the *only* ought-norms. To say that the ought-norms only bear on the trying seems to miscast our thinking about the norms of forgiveness. Naturally, one might reply by *identifying* the thing that is under our voluntary control with forgiveness. That way, our theory of forgiveness secures both the ought-norms and voluntary control. But as I noted above, taking this route is to give up on the RTF, the view that to forgive is to overcome resentment. This again reveals how difficult it is to articulate a version of the RTF that can account for both forgiveness ought-norms and our voluntary control over forgiveness.

A third and, to my mind, more formidable objection to this argument grants that ought norms do have a rightful place in our forgiveness practices, but rejects the claim that these norms presume that the victim can only comply with such norms via a voluntary exercise of her agency. Here, the thought is that when we invoke forgiveness ought-norms, we are not implying that the victim can forgive by, say, making a decision, forming a policy, or engaging in certain overt actions (e.g. an utterance of ‘I forgive you’). In other words, the critic can argue that ought-norms *do* bear on forgiveness itself, even though we *do not* have voluntary control over our forgiving.

To adopt such a view would be explicitly to defend a *non-voluntarist* account of forgiveness. Non-voluntarists about forgiveness claim that the fact that we lack voluntary control over forgiving does not mean that our forgiving cannot be subject to ought-norms; nor would it mean that we could not be morally responsible for our forgiving. As many non-voluntarists about moral responsibility in general have argued, it is commonplace and unproblematic to think that agents can be morally responsible and blameworthy for things that are not under their voluntary control.¹⁹ Presumably, some of these cases of responsibility and blameworthiness are for violating moral ought-norms. For example, Robert Adams claims that ‘[e]xaggerated or senseless anger [...] is morally offensive; and one who is guilty of it is liable to blame’ (1985, 3).²⁰ While it is true that one ought not manifest senseless anger, and also true that such a manifestation may not be under one’s voluntary control, this does not mean that one cannot be morally responsible and blameworthy for one’s senseless anger.

On this view, the exercises of our rational agency that are involved in revising THREAT are the kinds of things that are answerable to ought-norms, and they are also the kinds of things for which we can appropriately be held responsible were we to violate the relevant ought-norms. This is so even though none of the three stages of her account noted above picks out something that is under

our voluntary control. Rather, these judgments are under our control, in the sense that we are the ones making these judgments, and we are the ones who can revise them. Our judgments can reveal something about the kind of moral beings we are – how we assess and interpret the world and how we respond to moral reasons – and there are ways that we ought to do such things, morally speaking.

To further motivate this non-voluntarist line of thought regarding forgiveness, Hieronymi might point out that ought-norms do indeed bear on, for example, the making of racist judgments, the persistence of which one may not have much (if any) voluntary control over. But the fact that someone has little or no voluntary control over one's racist judgments does not entail that one *ought not* revise those judgments. So it is with one's judgment of THREAT: one might be obligated to revise (or not to revise) THREAT even if it is not in one's voluntary power to do so. But this does not mean that it is not under our power *at all* whether we revise THREAT. Revising THREAT is an exercise of our *rational* agency, and whether or not we revise our judgment that THREAT in the appropriate contexts says something about us – something about the kind of rational creatures we fundamentally are.

For now, I will sidestep a discussion of the voluntarism/non-voluntarism debate as it regards moral responsibility in general and focus on forgiveness in particular.²¹ There are two different issues here, and separating them can help us see how best to address this non-voluntarist way of understanding Hieronymi's account of forgiveness. One issue is whether we should favor a voluntarist or non-voluntarist account of forgiveness as such. The second issue is whether the ought-norm argument I have given actually supports the voluntarist position in light of the above non-voluntarist objection. We can take these issues in turn.

Consider first the general issue: Which sort of account of forgiveness – voluntarist or non-voluntarist – should we favor? While I think that *some* manifestations or modes of forgiveness may be non-voluntary or have important non-voluntary elements, I do not think that our general theory should rely on this way of thinking of forgiveness.²² Consider the fact that sometimes we ask others to forgive us. In making such requests, wrongdoers plausibly presume that the victim could voluntarily do something to grant the request. When asked to forgive, victims might respond with 'Yes, I forgive you,' or instead, 'I'm sorry, but it's too soon, and I don't want to forgive you yet.' There is a natural reading of these replies that suggests, minimally, that the victim *regards herself* as having voluntary control over her forgiveness. The practices surrounding our requesting forgiveness (and characteristic responses to such requests) therefore speak in favor of a generally voluntarist theory of forgiveness.²³

But now let us return to the question of whether the ought-norm argument for forgiveness voluntarism can withstand the non-voluntarist challenge. The non-voluntarist rebuttal points out that, in other contexts, ought-norms bear on things that are not under our voluntary control (such as racist judgments),

and so concludes that my argument does not support voluntarism after all. By way of reply, it is important to point out that not all oughts are created equal. Many philosophers have distinguished between what might be called an 'ought to be' and an 'ought to do' (see, e.g. Marcus 1966; Humberstone 1971, Harman 1977, Manne 2011; Schroeder 2011).

There are a number of ways to parse this distinction, but here I focus on the way Derk Pereboom has done so in recent work (2014). According to Pereboom, one kind of ought is the ought of *axiological evaluation* (or *axiological ideality*) (139). When employing these oughts, we make value-claims about states of affairs. And yet these oughts do not (at least directly) entail a 'can' claim (139). And so when we say, 'Bill ought to be more kind,' we are not implying that being more kind is something that Bill can do by voluntary agency (as opposed to, say, doing more kind actions). On the other hand, Pereboom notes another kind of ought: the ought of *specific action demand*. These oughts, in contrast to the first kind, 'do express a demand of an agent in a particular circumstance, which does entail that the agent can perform the indicated action' (139). And so when we say, 'Bill ought to talk to Jill more kindly,' we are implying that he can do something voluntarily.

With this distinction in mind, here is what I propose by way of reply to the non-voluntarist. As I see things, the oughts that bear on forgiveness are (at least in some cases) what Pereboom calls oughts of specific action demand. Such oughts require that a victim act (or not act) in certain ways and they presume that the victim can perform (or not perform) the relevant action (in this case, forgiving). Such oughts could *not* bear on forgiveness if we lacked voluntary control over our forgiving. And so since Hieronymi's account fails to give us voluntary control over our forgiving, it therefore cannot explain these ought norms.

It is worth asking whether Hieronymi's account of forgiveness, whatever its drawbacks, secures a kind of control over forgiveness that non-articulate accounts lack. Now, if by 'non-articulate' account, we just mean any account of forgiveness that does not implicate *any* kinds of revisions of judgment, then I think that articulate accounts will indeed secure a kind of rational control that non-articulate accounts lack. But it seems to me that non-articulate accounts could involve (or be premised upon) certain kinds of revisions to one's judgments without holding (1) that there is some *specific* judgment that is crucially implicated in forgiveness; or (2) that fundamentally, forgiveness just is a change in judgment that results in certain emotional changes.

Furthermore, my discussion may give the impression that some voluntary action or other is a necessary condition on forgiveness. Must one do *something* voluntarily (e.g. make a decision or commitment) in order to forgive? I suspect so, at least when it comes to the paradigmatic cases of forgiveness. Forgiveness is commonly thought to invoke what Marilyn McCord Adams called 'agent effort' (1991, 284). Why? Aside from the ought argument I have given above, I think there are a number of other reasons. I will only briefly mention them here.

First, forgiveness can be thought of as the manifestation of a moral virtue. And like other manifestations of moral virtue (courageous acts, honest acts, compassionate acts, etc.), forgiving acts will require voluntary exercises of agency. Second, forgiveness is paradigmatically a response to voluntary acts: requests for forgiveness, apology, acts of contrition, and the like. In responding to such acts, there is, I think, a natural expectation that one could (and typically would) do so by a voluntary exercise of agency. And third, in paradigmatic cases, forgiving alters the normative landscape. When I forgive you, I relinquish certain rights to blame and release you from certain personal obligations (e.g. to apologize, give further restitution, etc.) (Warmke [forthcoming](#)). This is why after I forgive you, I should no longer keep asking for apologies or censure you, for example. The best explanation for why forgiveness alters this normative landscape is due to a performative act on the victim's part that functions to give up rights to blame and release from personal obligations (Nelkin 2013; Warmke [forthcoming](#); Warmke [in preparation-b](#)).²⁴

At any rate, my conclusion, put simply, is this: some of the norms bearing on forgiveness are specific action demand ought norms. These norms presuppose that one's compliance is a matter of an exercise of voluntary agency. Hieronymi's non-voluntarist account of forgiveness therefore cannot explain these norms that bear on forgiving.

4. Forgiveness and obligation

We now turn to a second objection to Hieronymi's articulate account of forgiveness. To bring this objection into view, recall that, in developing her account, Hieronymi focuses on cases of forgiveness in which 'apology brings about a change in view or revision in judgment that allows one to forgo resentment' (2001, 545). Her account, then, should not be taken as an analysis of all cases of forgiveness, or even as an analysis of all cases in which one's forgiveness is principally a matter of eliminating resentment. But to the extent that Hieronymi's account is intended to capture the central or paradigmatic cases in which one forgives in response to a sincere apology (545), I will argue that her account has an undesirable consequence: it imbues wrongdoers with the power to *rationaly obligate* victims to forgive.

The basic idea, which I will develop in more detail below, is this: on Hieronymi's account, a wrongdoer can, by issuing an apology, retract the presumptive threat made by the past wrong action. In doing so, the wrongdoer can thereby undermine the victim's rationale for judging THREAT, the key judgment undergirding resentment, according to Hieronymi. And if the wrongdoer can do this, she can rationally obligate the victim to forgive her.

Here is the argument, put more precisely:

- (1) If a wrongdoer *W* issues a sincere apology to her victim *V* (who judges WRONG, RESPONSIBLE, SELF-WORTH, and THREAT), then *mutatis mutandis* *W* can do something that rationally undermines *V*'s judgment that THREAT (but not *V*'s judgments that WRONG, RESPONSIBLE, or SELF-WORTH).

Premise (1) follows from Hieronymi's account of how apologies typically work. Because an apology retracts the threat made by the wrongdoer's past action, then the victim's judgment that THREAT is no longer rationally founded, all other things being equal.²⁵

- (2) If *W* can do something that rationally undermines *V*'s judgment that THREAT (but not *V*'s judgments that WRONG, RESPONSIBLE, or SELF-WORTH), then *mutatis mutandis* *W* can do something that rationally obligates *V* to revise THREAT (but not WRONG, RESPONSIBLE, or SELF-WORTH).

Premise (2) follows from the fact that if the victim has been presented with sufficient evidence that THREAT is now no longer rationally founded, then the victim should, rationally speaking, give up her judgment that THREAT, assuming all other things are equal. By sincerely apologizing, the wrongdoer presents the victim with sufficient evidence that THREAT is no longer rationally founded. She has therefore done something that, rationally speaking, obligates the victim to give up her judgment that THREAT.

- (3) If *W* can do something that rationally obligates *V* to revise THREAT (but not *V*'s judgments that WRONG, RESPONSIBLE, or SELF-WORTH), then *mutatis mutandis* *W* can do something that rationally obligates *V* to forgive.

Premise (3) follows from Hieronymi's account of forgiveness: to revise THREAT (while keeping the other trinity of judgments) is (typically) to effect an elimination of resentment, which is what it is to forgive on her view. So a rational obligation to do the former is therefore a rational obligation to do the latter.

- (4) Therefore, if *W* issues a sincere apology to *V*, then *mutatis mutandis* *W* can do something that rationally obligates *V* to forgive.

This consequence, however, is undesirable for at least two reasons. To see the first reason, notice that, as Linda Radzik has pointed out, apologies and requests for forgiveness are distinct practices with distinct purposes (2009, 96). It is plausible to think that one may, without being confused or speaking infelicitously, say to one's victim immediately after a wrongdoing,

I'm so very sorry I did that, I apologize with all my heart. I am not asking you to forgive me, nor do I expect you to. Indeed, it may be the right thing for you to not forgive me at the point.

One can 'separate' oneself from one's past action and express regret and remorse without thereby thinking that forgiveness is something that the victim ought to do.

Yet it is difficult to make sense of this possibility if what one is doing when one apologizes is something that rationally obligates the victim to forgive. One would be doing something that (a) rationally undermines the victim's judgment that THREAT (and so rationally obligates forgiveness *per* the above argument), but also (b) expresses one's conviction that the victim need not forgive and perhaps even a desire that the victim not forgive at that time. This paints a portrait of wrongdoers who, upon expressing a very natural and reasonable sentiment, ultimately possess very divided psyches.²⁶

A second and perhaps more serious problem with the conclusion that wrongdoers can do something that rationally obligates a victim to forgive is that such a view simply gives wrongdoers too much power over the norms governing forgiveness. To bring this problem into view, notice that Hieronymi's account of forgiveness is committed to:

(SR) Sometimes, a sincere apology provides a requiring reason to forgive.

This is because, in her view, it is sometimes the case that a sincere apology retracts the relevant threat, and thereby undermines the victim's rationale for judging that THREAT. In the absence of sufficient countervailing reasons in favor of judging that THREAT, the victim is thereby rationally required to give up this judgment. A sincere apology can therefore be a requiring reason to forgive.²⁷ SR, however, is false, or so I will argue.²⁸

There are a number of common ways of speaking about forgiveness that suggest that SR is false. For example, forgiveness is commonly thought of as being a free gift. Therefore, to the extent that we cannot rationally require others to give us free gifts, then in like manner, wrongdoers cannot rationally require their victims to forgive them. Forgiveness is also commonly described as involving the canceling of a debt. Therefore, to the extent that debtors cannot rationally require that another cancel one's debt, then in like manner, wrongdoers cannot rationally require their victims to forgive them. Yet neither of these considerations constitutes much of an *argument* against SR. Moreover, each relies on an analogy that some reject.

So instead consider the following argument:

- (1) If SR is true, then wrongdoers, in virtue of doing wrong, sometimes gain the power to obligate victims to forgive them.

Premise (1) captures the fact that according to SR, wrongdoers have the ability to do something (i.e. apologize) that obligates their victim to forgive them. We can say therefore that their wrongdoing gives them a power to obligate their victim to forgive.

- (2) If wrongdoers, in virtue of doing wrong, sometimes gain the power to obligate victims to forgive them, then wrongdoers, in virtue of doing wrong, sometimes gain the power to obligate their victims to benefit them.

Premise (2) says that if you can obligate your victim to forgive you, you can obligate your victim to benefit you. The suppressed premise here is that to forgive someone is to benefit her. I will not defend this assumption in detail but will only say that it is natural to think that to be forgiven is to have something good done for you in one respect or another. I will leave unspecified the exact sorts of benefits at issue.

- (3) But wrongdoers do *not*, in virtue of doing wrong, sometimes gain the power to obligate their victims to benefit them.

Premise (3) is denial of the consequent of (2). It is hard to know what more could be said in defense of (3). It just seems obvious that doing wrong to another doesn't give you the power to obligate them to benefit you. Then, with two applications of *modus tollens*, we arrive at:

- (4) Therefore, SR is not true.

If this argument is sound, then a wrongdoer cannot obligate her victim to forgive her simply because the wrongdoer sincerely apologizes. But I think this is the result we want. It may be plausible that a sincere apology sometimes provides a *justifying* reason to forgive. It may even be plausible that a sincere apology always provides a justifying reason to forgive. But in my view, it is not plausible that a sincere apology that undermines THREAT can provide a requiring reason to forgive because this would mean that wrongdoers have the power to obligate their victims to benefit them.²⁹

Something has gone wrong, then, if Hieronymi's accounts of apology and forgiveness permit such obligating powers. But what exactly has gone wrong? On Hieronymi's account, to forgive is to overcome resentment by revising a certain sort of judgment about the world, namely the judgment that the wrongdoer's past action stands as a present threat. But the rationality of this judgment rests on a fact about the wrongdoer over which the wrongdoer *herself* has control. By apologizing, the wrongdoer can make it the case that the past action no longer stands as a present threat; the wrongdoer can 'cut off' the past wrong action from any continued meaning, thereby making continued anger pointless (2001, 548). And this worry generalizes to *any* view of forgiveness that (1) picks out a key judgment that must be revised in order to forgive, and (2) makes the rationality of that judgment dependent upon a fact that the wrongdoer can change. The lesson here, then, is that our theories of forgiveness should not permit wrongdoers to obligate their victims to forgive them, for doing so results in undesirable consequences about the norms of forgiving.

6. Normative constraints

I have argued that Hieronymi's account of forgiveness commits her to two implausible theses regarding the norms of forgiveness. First, her account entails that whether one forgives or not is largely outside the forgiver's voluntary control, and yet voluntarist accounts of forgiveness best explain the ought-norms that bear on forgiveness. Second, Hieronymi's account entails that wrongdoers can, by apologizing, rationally obligate their victims to forgive them. But because wrongdoers, in virtue of doing wrong, do not have the power to obligate their victims to benefit them, they therefore do not have the power to obligate their victims to forgive them.

These problems with Hieronymi's account, however, prove instructive: they make clear the extent to which one's commitments regarding the norms of forgiveness depend on, in some manner or another, one's commitments regarding the nature of forgiveness. To see this, consider the fact that most discussions of the norms bearing on forgiveness center on what we can call the *permissibility question*: Under what conditions (if any) is forgiveness morally permissible (or perhaps appropriate or fitting)? According to the *pessimist*, there are no conditions that either the wrongdoer or victim could meet that would make forgiveness as such morally permissible.³⁰ Most philosophers, however, are conditionalists of some sort, holding that, generally speaking, forgiveness is permissible only if certain kinds of conditions are met, either by the wrongdoer (e.g. Kolnai 1973, Haber 1991; Murphy 2003), or by the victim (e.g. Holmgren 1993; Garrard and McNaughton 2011), or by both the victim *and* wrongdoer (e.g. Griswold 2007). Unconditionalists, on the other hand, hold that there are no conditions that must be met either by the victim or by the wrongdoer in order for forgiveness as such to be permissible.³¹

What goes largely unnoticed in these discussions is the extent to which one's commitments regarding the *norms* of forgiveness depends upon, in some manner or another, one's commitments regarding the *nature* of forgiveness. For example, consider the rather simple example of two philosophers who agree that forgiveness is the overcoming of resentment. Yet while Philosopher A holds that resentment is a vindictive, hostile passion which aims to inflict suffering on another, Philosopher B holds that resentment is a natural and innocent sort of moral protest against wrongdoing. As a result, Philosopher A will be more likely to hold that we can have an obligation to forgive than Philosopher B. The basic – and I hope fairly obvious – point here is that what we say about the nature of forgiveness has consequences for what we say about the norms of forgiveness.

I therefore propose that one lesson to learn from our discussion of Hieronymi's account of forgiveness is that our theorizing about the nature of forgiveness should be informed and constrained by certain intuitive views about the norms bearing on forgiveness. In light of our above discussion, I suggest at least such two constraints: (1) that our theories of the nature of forgiveness should preserve

the full range of ought-norms that are typically thought to bear on acts of forgiveness; and (2) that our theories of the nature of forgiveness should not imbue wrongdoers with the power to obligate their victims to forgive them.

Naturally, there will be other plausible normative constraints on our theories of forgiveness. But let me suggest a modest methodological proposal: that our best theories of forgiveness should, so far as they are able, not undermine certain intuitive views about the norms bearing on forgiveness. To be clear, I am not claiming that our intuitive views about the norms of forgiveness are unassailable. Perhaps some of our commonsense views about the morality of forgiveness are mistaken. Nor do I mean to imply that there is a clear, monolithic, and widely agreed-upon set of norms bearing on forgiveness. However, I do claim that if a theory about the nature of forgiveness has the implication that certain widely held assumptions about the norms of forgiveness are false, then that is a significant strike against that theory. I am therefore advocating for a kind of reflective equilibrium for theorizing the nature of forgiveness, one that would not only take into consideration things such as our best evidence from the psychological sciences, but also common views about the morality of forgiveness.

Why use normative constraints when theorizing the nature of forgiveness? I conclude by offering two considerations for further consideration. First, reflection on the norms bearing on forgiveness can offer important insight into the *kind* of phenomenon that forgiveness is. For example, if it is widely thought that forgiveness sometimes answers to specific action ought-norms, then that is good evidence that whatever phenomenon our theory targets as forgiveness should be the kind of thing that answers to ought-norms. The evidence is defeasible, of course. On reflection, we may come to see that our norms need altering. But jettisoning these norms should only come after hard-won philosophical toil. An account of the nature of forgiveness should not come out of the gate assuming that certain common sense assumptions about the ethics of forgiveness are mistaken.

Second, and relatedly, we sometimes have a better grasp of the norms bearing on some moral practice than we do the particulars of the nature of the practice itself. I think this is true of forgiveness and I think it is true of philosophers and laypersons alike. Studies of lay views on forgiveness suggest a very general and very diverse set of views.³² While there is surely plenty of disagreement about the norms bearing on forgiveness, there is also, I hazard a guess, general consensus, that, for example, sometimes one ought (or ought not) to forgive, or that, in general, wrongdoers do not have the prerogative to obligate their victims to forgive them.³³

I suspect that something like this is true for philosophers as well. Do we, as philosophers, have good reason to be more confident in holding that forgiveness crucially implicates the revision of a very specific judgment (i.e. THREAT) than we are in holding that forgiveness sometimes answers to specific action

ought-norms or that wrongdoers generally cannot obligate their victims to forgive them? If not, this is a good evidence for utilizing normative constraints in constructing theories of the nature of forgiveness.³⁴

Notes

1. A bit of terminology will help us proceed. I will use the term 'wrongdoer' to refer to an agent who is a putative candidate for being forgiven, and the terms 'wrong' and 'wrongdoing' to refer to the conduct for which a wrongdoer is forgiven. I will also use the term 'victim' to refer to an agent who is putatively in a position to forgive a wrongdoer. I intend for nothing more to be inferred from the use of these terms.
2. See Kekes (2009, 490), Radzik (2009, 117), and Zaibert (2009, 38).
3. Bash (2007, 161).
4. See Richards (1988, 184) and Darwall (2006, 72), respectively.
5. See Griswold (2007, 40) and Lauritzen (1987, 142), respectively.
6. Nelkin (2013) rejects both the RTF as a general claim about forgiveness as well as Hieronymi's specific account of how we forgive by overcoming resentment. Zaragoza, however, agrees with Hieronymi that forgiveness crucially implicates the overcoming of resentment and that our best accounts should be both 'articulate' and 'uncompromising' (notions we shall unpack shortly) (2012, 604). Zaragoza's disagreement, however, is with Hieronymi's particular account of how resentment is overcome when one forgives. I will not pursue his objections to her account here, for my concern is what consequences her view has as regards the norms bearing on forgiveness. So far as I can tell, his revisions to her account do not avoid the kinds of criticisms that I raise here.
7. Hieronymi's primary target here is David Novitz, who claims that '[e]ven if there is no act of forgiving that can directly terminate my resentment and rage, there is nonetheless a good deal that I can do indirectly in order to ease my negative feelings [...] Any such attempt to identify imaginatively with your situation, and the resultant (perhaps imperfect) understanding of what it was like to be in that situation, may help destabilize my attitudes towards you' (1998, 309).
8. In this passage, Hieronymi appears to be endorsing what we can call an *objective account* of forgiveness, which says that in order to forgive, one must *actually* be the object of blameworthy treatment. This is why forgiveness involves the overcoming of 'justified' resentment. *Subjective accounts*, on the other hand, would claim that one could forgive by overcoming resentment, even if that resentment was unjustified (perhaps because one was not actually the victim of blameworthy treatment). I leave this matter unsettled.
9. Why does Hieronymi think that articulate accounts are superior? It is hard to say exactly, and she herself claims to give no direct arguments for such accounts (2001, 536). However, she does suggest the following sort of motivation: '[If] both resentment and forgiveness admit of justification, i.e. if one resents or forgives another person thinking one has a good reason to, then forgiving will entail more than figuring out how to rid oneself of certain unfortunate affects. Ridding oneself of resentment by taking a specially-designed pill, for example, would not count as forgiveness. Genuine forgiveness must involve some revision in judgment or change in view. An account of forgiveness must therefore *articulate* that revision in judgment or change in view. It must be an articulate account' (2001, 530).
10. See, for example, Murphy (2003, 13–15) on this point.

11. Two clarifications are in order. First, Hieronymi is careful to note that resentment involves more than merely assenting to the truth of the relevant judgments, but additionally involves some sort of *commitment* to them (530). She describes the difference as being the same sort of difference between 'agreeing that something is good and wanting it, or agreeing that something is dangerous and fearing it' (530). However, I shall follow Hieronymi and speak simply in terms of the making and revising the relevant judgments. Second, Hieronymi does not claim that her account captures every case of forgiveness: she limits her analysis to cases of forgiveness in which 'an apology brings about a change in view or revision in judgment that allows one to forgo resentment' (2001, 545). She does, however, think that her account is the sort that is 'required' (545): any theory must be articulate and uncompromising, identifying a unique judgment that is rationally revised when one forgives.
12. In his careful 2012 discussion of Hieronymi's paper, Zaragoza formulated and labeled the judgments that Hieronymi claims are constitutive of resentment. I borrow those formulations and labels here.
13. Cf. Murphy and Hampton (1988, 25): moral injuries 'are *messages* – symbolic communications. They are ways a wrongdoer has of saying to us, "I count but you do not," "I can use you for my purposes," or "I am here up high and you are there down below."'
14. It is worth noting that Hieronymi holds that the object of resentment is 'the *claim* made by the event rather than the *author* of the event' (2001, 547). Here, Hieronymi stands in disagreement with the mainstream of literature on forgiveness. Typically, authors have claimed that the target of resentment is either the action (e.g. Murphy and Hampton 1988, 60) or the author of the action (e.g. Griswold 2007, 24).
15. See (2001, 535).
16. Of course, there are things that we can do voluntarily that are closely related to justifying and excusing. For example, we may *decide* no longer to treat someone as blameworthy, or we may *try* to bring about these changes in our orientation to the other person, or we may initiate or resume friendly relations. But as I understand the matter, these are not the same as justifying or excusing.
17. Novitz (1998) briefly noted in passing this connection between the ought-norms bearing on forgiveness and control. While Novitz himself rejects the view that forgiveness is sometimes under our voluntary control (313), he claims that to hold the view that we sometimes ought to forgive implies 'the view that people can choose to forgive, and can somehow forgive at will' (299–300, fn. 1). In other words, Novitz would accept my premise (2) in principle, but would reject the view that we have or need such control over forgiving. What are Novitz's reasons for rejecting the view that we sometimes possess voluntary control over our forgiveness? As best I can tell, it is simply due to a commitment to the view that to forgive is to overcome an emotion, and that this is not something that we can do by a voluntary exercise of agency (313). While I agree with Novitz that such emotional change is not under our voluntary control, I take this as a reason to reject such theories of forgiveness, whereas he, it seems, does not.
18. Michael McKenna and I suggested (but did not fully develop) such an argument in our 2013 paper.
19. Non-voluntarists reject the voluntarist thesis, which is the view that, as Michael McKenna puts it, 'A person is morally responsible only for those things that are within the control she is able to exercise by acting freely' (2012, 187).

20. McKenna helpfully lists a number of things that non-voluntarists have claimed can serve as grounds for blame. He notes that Adams (1985, 18) considers 'believing that certain people do not have rights that they have; perceiving members of some social group as less capable than they actually are; failing to notice indications of other people's feelings; and holding too high an opinion of one's attainments.' Similarly McKenna (2012, 190) points out that on Angela Smith's non-voluntarist position, 'one can be blameworthy for forgetting a friend's birthday (2005, 236), finding amusing something that is morally offensive (237), being emotionally indifferent to intimates (242–243), and having morally objectionable thoughts occur to us, even if we elect not to act upon them (246).'
21. For an overview of the debate and a proposed compromise between these two positions see McKenna (2012, 187–195).
22. I affirm the possibility of a wholly 'private' mode of forgiveness that is constituted by some constellation of belief revisions and emotional changes, and perhaps also certain kinds of decisions or policy-formations (see, e.g. Warmke in preparation-a). But as I have been at pains to show (and for further reasons I will elucidate below), I think such cases are not the paradigmatic or central cases of forgiveness and so should not be the basis of a theory of the nature or norms of forgiveness.
23. There is another issue looming here that I will mention only briefly. If you recall, Hieronymi holds that her articulate account secures the forgiver more control over her forgiveness than do non-articulate accounts. But consider a case of non-articulate forgiveness in which one overcomes resentment, not by rationally revising a singular judgment such as THREAT, but rather by: (1) voluntarily driving to meet with a therapist once a week to discuss the past wrong and ways to move past it; and (2) the disappearance of resentment over time (in a way that does not collapse the forgiveness into excuse, justification, or condonation). We can assume in this case that the resentment is not eliminated by rational revision of the sort that Hieronymi has in mind. And yet the disappearance of resentment is distally brought about by a series of deliberate, voluntary free actions that the victim undertakes in order to accomplish forgiveness. Now we can press the question: What is the principle according to which this non-articulate account of forgiveness secures the forgiver *less control* over her forgiveness than would an articulate account whereby forgiveness is accomplished non-voluntarily? The point is that once we lay bare the fact that Hieronymi's non-articulate forgiveness is accomplished by non-voluntary exercises of agency, it becomes less clear that such an account is superior to non-articulate accounts with respect to the control it secures the forgiver.
24. I thank an editor for this journal for raising these points.
25. Presumably, there will be cases in which even though a wrongdoer issues a sincere apology and retracts the presumptive THREAT made by the past wrong action, the victim will have other kinds of evidence that justifies skepticism that the threat really has been retracted. Perhaps, for example, the victim knows that in past situations, the wrongdoer has insincerely apologized and so has not really retracted the threat. This might justify skepticism even in those cases in which the wrongdoer really does retract the threat. To simplify matters, I will proceed by considering only those cases in which the wrongdoer sincerely apologizes and retracts the threat, the victim correctly believes that this is what has occurred, and no other relevant factors are present. Thanks to Michael McKenna for raising this issue.
26. It might be objected that the wrongdoer isn't typically *aware* that she is rationally obligating the victim to forgive when she apologizes and is therefore

not committed to this sort of divided psyche. But this response implies that wrongdoers who apologize to, but do not want or expect forgiveness from their victims are confused about how their apology is functioning: they think they are expressing one's conviction that the victim need not forgive, yet they are doing something that effects a rational obligation to forgive.

27. Requiring reasons are often contrasted with *justifying* reasons. Roughly put, a reason justifies a course of action if it can provide a justification for a course of action that would otherwise be irrational. So, for example, it is irrational for me to jump in front of a moving bus. But suppose that, in doing so I could save the life of a stranger. This fact provides a justifying reason to jump in front of the bus (and it may do so even if I am not all things considered justified in jumping in front of the bus). A requiring reason would rationally require one to jump out in front of the bus. If the person I could save was instead my toddler, I may have a requiring reason to do so. For more on the justifying/requiring reason distinction, see Gert (2000).
28. Here, we should be careful to separate three different worries one might have about SR as it pertains to Hieronymi's account of forgiveness. First, one might object to SR on the grounds that it wrongly claims that apologies can rationally require victims to *revise* THREAT. Second, one might object to SR on the grounds that it wrongly claims that apologies can rationally require victims to *relinquish* their *resentment* (whatever our theory of resentment happens to be). Third, one might object to SR on the grounds that it wrongly claims that apologies can rationally require victims to *forgive* (whatever our theory of forgiveness happens to be). While it is true that on Hieronymi's view to do the first of these three things (i.e. revise THREAT) is typically also to do the latter two things (i.e. overcome resentment and forgive), one might think that even if Hieronymi is correct that apologies can rationally require a revision to THREAT, this does not entail that apologies can rationally require either the elimination of resentment or forgiveness. One might, for example, reject the claim that to resent requires first judging then revising THREAT, and so if one is rationally required to revise THREAT, this does not entail a rational obligation to give up resentment. Alternatively, one may reject the claim that to forgive requires first experiencing then giving up resentment, and so even if one can be rationally required to revise THREAT due to an apology and it is true that to rationally revise THREAT is to eliminate resentment, this would not entail that one can thereby be rationally required to forgive due to an apology. All that said, I simply want to argue that whatever else, SR is false because wrongdoers cannot effect an obligation that others forgive them.
29. I should be clear that I have not argued that there is never an all things considered obligation to forgive. Nor have I argued that wrongdoers could never do things that would make it the case the victims have an all things considered obligation to forgive them. I could threaten your child's life if you do not forgive me and thereby (we might presume) give you an all things considered obligation to forgive me. What I have argued, however, is that we should reject theories of forgiveness that imbue wrongdoers with the power to obligate their victims to forgive them by, say, apologizing to them: apologies should not be thought of as providing requiring reasons to forgive. Forgiveness is commonly thought of as a kind of gift or grace and there is something in this thought that strikes me as worth keeping. If wrongdoers could obligate their victims to forgive them by simply apologizing, then we have left the economy of gift or grace and entered the economy of desert and obligation.

30. Nietzsche might be thought of as a forgiveness pessimist, though the matter is complicated. See, e.g. Griswold (2007), and Blustein (2014).
31. I am unaware of anyone in the philosophical literature who defends such a view.
32. e.g. Kelley asked participants to write forgiveness narratives describing situations in which they forgave, were forgiven, or requested forgiveness. Participants described multiple types of forgiveness, including intellectual/emotional, conditional, and unilateral/bilateral forgiveness [...] Laypersons reported thinking about and choosing forgiveness, but they also discussed forgiveness as a change in feelings towards the transgressor' (as discussed in Waldron and Kelley 2008, 13).
33. Whether these views are correct is, of course, another matter.
34. I am grateful to Michael Gill, Michael McKenna, and Mark Timmons for their careful feedback on earlier versions of this paper. I also thank two referees and an editor from this journal for the time and attention they gave to my ideas. Support for this work was funded in part by a grant from the Templeton World Charity Foundation. The opinions expressed in this paper are my own and do not necessarily reflect the views of the foundation.

Notes on contributor

Brandon Warmke is a postdoctoral research associate at the University of Notre Dame. He primarily works in ethics and moral psychology and has published a number of articles on forgiveness, punishment, and moral responsibility.

References

- Adams, Robert Merrihew. 1985. "Involuntary Sins." *The Philosophical Review* 94: 3–31.
- Adams, Marilyn Mc Cord. 1991. "Forgiveness: A Christian Model." *Faith and Philosophy* 8 (3): 277–304.
- Bash, Anthony. 2007. *Forgiveness and Christian Ethics*. Cambridge: Cambridge University Press.
- Blustein, Jeffrey M. 2014. *Forgiveness and Remembrance: Remembering Wrongdoing in Personal and Public Life*. New York : Oxford University Press.
- Calhoun, Cheshire. 1992. "Changing One's Heart." *Ethics* 103: 76–96.
- Darwall, Stephen. 2006. *The Second Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Garrard, Eve, and David McNaughton. 2002. "In Defence of Unconditional Forgiveness." *Proceedings of the Aristotelian Society* 103 (1): 39–60.
- Garrard, Eve, and David McNaughton. 2011. "Conditional Unconditional Forgiveness." In *The Ethics of Forgiveness*, edited by Christel Fricke, 97–106. New York: Routledge.
- Gert, Joshua. 2000. "Practical Rationality, Morality, and Purely Justificatory Reasons." *American Philosophical Quarterly* 3: 227–243.
- Griswold, Charles L. 2007. *Forgiveness: A Philosophical Exploration*. New York: Cambridge University Press.
- Haber, Joram Graf. 1991. *Forgiveness*. Lanham, MD: Rowman and Littlefield.
- Harman, Gilbert. 1977. *The Nature of Morality*. Princeton, NJ: Princeton University Press.
- Hieronimi, Pamela. 2001. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62 (3): 529–555.

- Holmgren, Margaret. 1993. "Forgiveness and the Intrinsic Value of Persons." *American Philosophical Quarterly* 30 (4): 341–352
- Holmgren, Margaret. 2012. *Forgiveness and Retribution*. Cambridge: Cambridge University Press.
- Hughes, Paul M. 1993. "What is Involved in Forgiving?" *The Journal of Value Inquiry* 27: 331–340.
- Humberstone, I. L. 1971. "Two Sorts of Oughts." *Analysis* 32: 8–11.
- Kekes, John. 2009. "Blame versus Forgiveness." *The Monist* 92: 488–506.
- Kolnai, Aurel. 1973–74. "Forgiveness." *Proceedings of the Aristotelian Society* 74: 91–106.
- Lauritzen, Paul. 1987. "Forgiveness: Moral Prerogative or Religious Duty?" *Journal of Religious Ethics* 15 (2): 141–154.
- Manne, Kate. 2011. "Not by Reasons Alone." PhD diss., Cambridge, MA: MIT Press.
- Marcus, Ruth Barcan. 1966. "Iterated Deontic Moralities." *Mind* 75: 580–582.
- McKenna, Michael. 2012. *Conversation and Responsibility*. Oxford: Oxford University Press.
- Murphy, Jeffrie G. 2003. *Getting Even: Forgiveness and Its Limits*. New York: Oxford University Press.
- Murphy, Jeffrie G., and Jean Hampton. 1988. *Forgiveness and Mercy*. Cambridge: Cambridge University Press.
- Nelkin, Dana. 2013. "Freedom and Forgiveness." In *Free Will and Moral Responsibility*, edited by Ishtiyaque Haji and Justin Caouette, 165–188. Newcastle upon Tyne: Cambridge Scholars Press.
- Novitz, David. 1998. "Forgiveness and Self-respect." *Philosophy and Phenomenological Research* 58: 299–315.
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Radzik, Linda. 2009. *Making Amends: Atonement in Morality, Law, and Politics*. Oxford: Oxford University Press.
- Richards, Norvin. 1988. "Forgiveness." *Ethics* 99: 77–97.
- Schroeder, Mark. 2011. "Ought, Agents, and Actions." *Philosophical Review* 120: 1–41.
- Smith, Angela. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236–271.
- Strawson, P. F. 1962. "Freedom and Resentment." In *Free Will: Second Edition*, edited by Gary Watson, 72–93. Oxford: Oxford University Press, 2003.
- Waldron, V. L., and D. L. Kelley. 2008. *Communicating Forgiveness*. Thousand Oaks, CA: Sage.
- Warmke, Brandon, and Michael McKenna. 2013. "Moral Responsibility, Forgiveness, and Conversation." In *Free Will and Moral Responsibility*, edited by Ishtiyaque Haji and Justin Caouette, 189–212. Newcastle upon Tyne: Cambridge Scholars Press.
- Warmke, Brandon. **Forthcoming**. "The Economic Model of Forgiveness." *Pacific Philosophical Quarterly*.
- Warmke, Brandon. **In preparation-a** "Modes of Forgiveness."
- Warmke, Brandon. **In preparation-b** "The Normative Significance of Forgiveness."
- Zaibert, Leo. 2009. "The Paradox of Forgiveness." *Journal of Moral Philosophy* 6 (3): 365–393.
- Zaragoza, Kevin. 2012. "Forgiveness and Standing." *Philosophy and Phenomenological Research* 84 (3): 604–621.