Punishment and Forgiveness

Forthcoming in *Routledge Handbook of Criminal Justice Ethics*
Jonathan Jacobs and Jonathan Jackson (eds)

Justin Tosi
University of Michigan

Brandon Warmke
University of Notre Dame

1. Introduction

There is, we suspect, a widely held assumption that forgiveness and punishment are crucially linked. Our sense is that many people think that if a victim forgives her wrongdoer then the victim gives up her right to punish the wrongdoer, or at least that if the victim *were* to punish someone she had forgiven, then she would have done something morally inappropriate. Such a view likely seems intuitively correct to many. We will argue that upon inspection it is not so appealing.

A number of philosophers have claimed there is an important link between punishment and forgiveness. In *Leviathan*, for example, Hobbes drew a connection between forgiving and the facility to pardon in his sixth Law of Nature: "A sixth law of Nature is this, 'that, upon caution of the future time, a man ought to pardon the offences past of them that, repenting, desire it'" (1969 [1651]). Remarking on this law, Bernard Gert writes, "This virtue, which Hobbes calls having the facility to pardon, one can also call being forgiving" (2010, p. 98). In a similar vein, Leo Zaibert has recently argued that to forgive is "deliberately to refuse to punish" (2009, p. 368). David Londey tells us that that a "clear" condition on forgiveness is that "in forgiving you I remit *any* penalty or sanction that your wrongdoing would otherwise bring on your head" (1986, pp. 4-5, italics original).[1] Nicholas Wolterstorff writes that one of the "main components" of forgiveness "is the foregoing of retributive punishment" (2009, p. 203). Relatedly, Richard Swinburne has claimed that "if I forgive you for some act, I ought not subsequently to punish you for that act" (1989, p. 87 fn. 8). Psychologists have also linked forgiveness and punishment. According to Robert Enright and his colleagues, forgiveness involves "the casting off of deserved punishments" (1992, p. 88). On these sorts of *punishment-forbearance* views of forgiveness, forgiving crucially implicates, in one way or another, the forbearance of punishment.[2]

---

[1] Londey clarifies his use of the phrase "penalty or sanction" to "include moral censure as well as more tangible punishments" (1986, p. 5).

[2] Claims like these just adumbrated are usually made in passing and are not given defense. One exception is Zaibert (2009). Because one of us has discussed that view elsewhere (Warmke 2011), we shall not pursue it here. We shall instead be concerned with what kinds of defenses *could* be marshaled for these claims, and with how plausible those arguments could be.

If these theorists are right about forgiveness, there could be important consequences for criminal justice ethics. Consider the following passage from Anthony Bash:

> Suppose W, a wrongdoer, defames V, a victim. The wrongdoing is actionable in the courts. Rather than go to court, V chooses to forgive W, because W and V are long-standing friends and because W is sorry about what she has done. By forgiving W, V implicitly declares (among other things) that W has wronged her and that she is setting aside the right to take legal proceedings against W for the wrongdoing. Legally, V can still sue W for the wrong (as her legal rights are unchanged); however, as I have said [earlier], it is contrary to what is implicit in what it means to forgive if V pursues her legal rights against W (2015, p. 53).

Suppose that Bash is correct that there is such a relationship between forgiveness and punishment: if one forgives, one should not press one's legal rights against and seek punishment of one's wrongdoer. This claim raises a number of practical issues. For example, if a wrongdoer has been forgiven by her victim, should this be taken into account when it comes to questions of what counts as a just punishment? If it was widely believed that if one forgives a wrongdoer, the wrongdoer should not be punished, would this have deleterious effects on the justice system? Would victims, for example, be less willing to testify in court or file charges if they believed that doing so is wrong because they have forgiven? Can states that are in a position to punish criminals also "forgive" them? And more generally, what relationship should forgiveness have to criminal justice systems in the first place?

In this paper we shall not attempt to answer all of these questions. We will, however, provide a basic philosophical framework for thinking about the moral relationship between forgiveness and punishment.[3] In Section 2, we will develop and motivate the kind of punishment-forbearance view we target in this paper: the claim that if one forgives a wrongdoer for some bit of conduct, one should not subsequently punish her for it. In Section 3, we turn to critically assess the arguments for this view. We conclude that none succeeds and that we have good reason to deny that forgiving a wrongdoer must make punishing her morally inappropriate. In Section 4 we turn to the issue of institutional forgiveness: can entities like states forgive individuals? We argue that to the extent that this is possible at all, such "forgiveness" will largely be a matter of simply issuing a pardon. We conclude in Section 5 by reflecting on the differences between forgiveness and acts of mercy.

---

[3] While punishment-forbearance views of forgiveness have not received systematic treatment, discussions of forgiveness and punishment can be found in Mabbott (1939), Morris (1968), Duff (1986, pp. 216-17), Haber (1991), Griswold (2007, pp. 32-3), Murphy (2003, p. 101), O'Shaughnessy (1967), Swinburne (1989), Garrard and McNaughton (2002), Zaibert (2009), Warmke (2011, 2013), Pettigrove (2012), and Jacobs (this volume).

2. Punishment and Forgiveness

In this section we explain and motivate one way of thinking about how forgiveness and punishment are crucially linked. Specifically, we will be exploring what kind of relationship forgiveness has to the *forbearance* of punishment. By "forbearance of punishment" we have in mind two things. First, to forbear punishing a wrongdoer is to commit not to punish her or to deliberately refuse to punish her. Second, by forbearing punishment, one also actually abstains from punishing. To forbear punishing a wrongdoer, then, is both to deliberately refuse to do so and then actually to refrain from punishment. This way of putting the matter makes clear that the victim who forbears punishment does so purposefully. She does not just, say, forget to punish; rather, her not punishing the wrongdoer is intentional. This also makes clear that one's forbearing punishment is not *just* to commit not to or refuse to punish; it is also to abstain from punishing the wrongdoer. Naturally enough, there are numerous other ways of exploring the relationship between punishment and forgiveness, but for now, we shall simply focus on the possible moral connections between forgiveness and the forbearance of punishment. We will suggest a few ways of arguing for some interpretation of the claim that forgiveness requires forbearance from punishment.

To begin, consider just one way of construing the relationship between forgiveness and the forbearing of punishment. According to this view, the relationship is a moral one, not a metaphysical or conceptual one.[4] In other words, this view makes no claims about, for example, whether forgiving *itself* involves forbearing punishment. Rather it makes a moral claim: forgiving a wrongdoer makes it morally inappropriate to punish her, and so if one forgives, one should forbear punishment. We might put the claim this way:

> If agent A forgives agent B for B's conduct X, then *ceteris paribus* A should forbear punishing B for X.

What would motivate such a view? Why might forgiving a wrongdoer make it morally wrong or inappropriate to punish her? Here, we explore three kinds of motivations.[5]

First, consider the fact that it is widely held that forgiving requires one to forswear both retributive attitudes like resentment as well as retributive behaviors. From this view, it is not a great leap to the thought that when one forgives one forfeits any right to visit

[4] Here are two other ways of thinking about the relationship between punishment and forgiveness. One is a necessity claim: Agent A forgives agent B for B's conduct X only if A forbears punishing B for B's having done X. The other is a sufficiency claim: Agent A forgives agent B for B's conduct X if A forbears punishing B for B's having done X. These statements make a metaphysical or conceptual claim about the relationship, and not, in the first instance, a moral claim. We shall not be addressing these other claims in this paper.

[5] While, as we noted above, it is sometimes claimed that punishment is linked to forgiveness, it is a claim that is rarely argued for. Hence we shall try to construct three plausible motivations for certain kinds of punishment-forbearance views.

unpleasant consequences upon the wrongdoer on the basis of what she did. After all, the right to make life unpleasant for the wrongdoer seems no longer to be something the forgiving agent will want to stand on. Punishment is (minimally) the visiting of unpleasant consequences in response to a wrong.[6] So, forfeiting the right to punish is simply a part of forgiveness. Well, quite obviously, once one has forfeited the right to punish a wrongdoer, one cannot rightly punish that wrongdoer for what she has done. To insist that one retains the right to punish in this case is simply to misunderstand what forgiveness involves. Therefore, to punish a wrongdoer when one has forgiven her would be wrong—one should forbear punishment instead. Call this the *forfeiture argument*.

Second, one plausible way of understanding forgiveness is as an abandonment of certain reasons for action. By abandonment we mean something like a dismissal or disavowal of the justifying force of those reasons. Once I abandon a reason, I can no longer appeal to it as a justification for my actions. Indeed, this is the very sort of thing that many think happens in the case of forgiveness and blame. When we forgive, we still judge that the wrongdoer is *blameworthy* and that we could have the *right* to blame the wrongdoer. And yet when we forgive, we forswear blame—we commit to not blame, we abandon justifying reasons to blame.

So perhaps when we forgive, we abandon all justifying reasons for punishment, and so to punish anyway would be inappropriate. Note that this argument does not depend on the success of the forfeiture argument. One may retain the right to punish and still have no good reason to do so. To punish without a good reason for doing so—a justifying aim—would be truly perverse.[7] Since punishment imposes costs, and often significant ones, on another person, it is not something we should do without a morally important goal in mind. But if forgiveness involves abandoning the reasons that could serve as justifying aims of punishment, then what we are left with in punishment is pointless and cruel treatment of another. Therefore, forgiving has the consequence that (barring exceptional circumstances) one is not permitted or justified in punishing someone that one has forgiven.[8] One should instead forbear punishment. Call this the *no justification argument*.

Third, it might also be argued that when we forgive, we enter into a new kind of relationship with the wrongdoer. We are to some degree reconciled and typically seek still further reconciliation. We welcome the wrongdoer back into the moral community. We endeavor to treat and regard her with love and good will, seeking to repair the

---

[6] The definition of punishment is an extremely controversial topic. This brief statement is arguably not sufficiently ecumenical, but it is roughly accurate, and should be sufficient for our purposes. It is loosely adapted from Hart (1968, pp. 4-6). For an extensive recent discussion of the definition of legal punishment, see Boonin (2008, pp. 3-28).

[7] We do not mean to claim here that the justification of punishment must be consequentialist. The term 'justifying aim' is borrowed from Hart (1968, pp. 8-11), and may refer more broadly to the purpose of the act of punishing itself rather than merely to its consequences.

[8] Why the caveat? It may turn out that in exceptional circumstances, even though it would be wrong for a victim to punish someone she has forgiven (as such), the balance of reasons might weigh strongly in favor of punishing anyway. If punishing is the only way to save a village of children, then it may be all things considered permissible to punish even if there are strong pro tanto reasons against punishing.

relationship as much as possible. In light of this new orientation, however, punishment will be out of place. If you are seeking to *repair* a relationship—to put it back on something approximating equal moral footing—then punishing the wrongdoer is manifestly *not* the way to go about doing so. On this sort of view, punishment is morally inappropriate given the kind of relationship that one has to the wrongdoer when one forgives her. It may be perfectly appropriate to punish the wrongdoer *before* one has forgiven, one might argue. But once you have forgiven, you've set the course for the relationship going forward, a course that ought not involve punishing the wrongdoer for the wrongdoing in question. Again, you should forbear punishment. Call this the *new relationship argument.*

3. Some Objections

Before treating the individual arguments canvassed above, we should note some general worries about punishment-forbearance views of forgiveness. First, usually when we are wronged, we are not at liberty to punish our wrongdoers. We might be able to blame them, or take revenge upon them, but neither of those is the same thing as punishment (see, e.g., McKenna 2012, Warmke 2013). We generally think of punishment as something institutions do—particularly, the state. Interpersonal punishment seems to be restricted to personal hierarchical relationships, as between a parent and child or a teacher and a student. Some political philosophers think that there is a natural right to punish wrongdoers, held by all, but those who hold this view generally think that we, as members of civil society, transfer it to the state upon leaving the state of nature.[9] We suspect, then, that talk of interpersonal punishment between equals within civil society will probably strike many readers as odd. So in order to get this view off the ground, we will suppose that punishment-forbearance theorists must take a wide view of what counts as punishment, so that it includes informal social sanctions in response to wrongdoing. Thus, ostracizing a wrongdoer, kicking her out of your home, breaking off personal ties, or perhaps even spreading the news of her wrongdoing far and wide—all of these acts would (or at least could) count as punishment.[10] Without a very inclusive conception of interpersonal punishment, we think it is very difficult even to get punishment-forbearance views off the ground.[11] For the purposes of this paper, then, we shall adopt a liberal view of what can count as punishment.

---

[9] Cf. Wellman (2009), though see Simmons (1991).

[10] Does forbearing punishment require one to "undo" one's previous punishment (or perhaps blaming) of her? Suppose you punish your child for smoking dope in his bedroom by kicking him out of the house. Does forbearing punishment require inviting him back into the house? Is that "undoing" his punishment, or ceasing his punishment? As best we can tell there is no good general answer to these questions, and they must be considered on a case-by-case basis.

[11] Strictly speaking, one could perhaps consistently claim that we forbear punishment when we forgive (or that we must not punish if we forgive), *even if* we are rarely if ever able to punish (because we lack standing to do so). That we cannot do something, one might argue, does not show that we cannot forbear doing it. Even granting this point, the view would be a vacuous one, and not worth discussing. We will assume then, that whatever view of punishment is at issue, it must be something that could be at issue in common cases of interpersonal forgiveness.

Second, we think that many conflate punishment with related but importantly different moral phenomena. For instance, we suspect that when many people use the word 'punishment' in discussions of forgiveness, they have in mind either something like overt moral blame (e.g., certain forms of censure, denunciation, reproach, withdrawal of friendly relations)[12] or instead revenge. However, we distinguish between all of these: these phenomena come apart, both conceptually and in practice. Now if defenders of the forbearance view thought that what they call punishment really just is (what we would call) overt blame or revenge, then we could agree that forgiveness is crucially linked to those phenomena: forgiveness typically does rule out overt blaming and revenge, and the forgiveness literature is in general agreement about this.[13] If by "punishment" the defender of this view just meant, say, overt blaming behaviors, then we will be the among the first to agree that forgiveness typically involves forbearing such reactions to the wrongdoer and also makes such reactions to the wrongdoer morally inappropriate. But thinking (as we do) that forgiveness typically rules out blame and revenge does not vindicate the defender of the punishment-forbearance view unless she *just equates* interpersonal punishment with either overt blame or revenge. In what follows, we will assume that the defender of punishment-forbearance views *does* think that punishment is something *distinct* from overt moral blame and revenge, and that they are not simply making a claim about a phenomenon of the forbearance of blame or revenge.

How might we distinguish between overt blame, revenge, and punishment? If forgiveness typically rules out blame and revenge, but not punishment, how would one go about showing this? One strategy would be to give full accounts of each of these phenomena. That is obviously beyond the scope of this paper or any paper. Another strategy would be to articulate certain key differences between the phenomena without giving full accounts. That way, we could differentiate between phenomena by asking whether they have certain key features. One of us (Warmke 2013) attempted this strategy elsewhere, and we shall not rehearse those arguments here. We shall pursue a different tack. Instead of trying to distinguish these phenomena, we shall simply argue that forgiveness does not have the tight relationship to forbearing punishment that punishment-forbearance

[12] Though a full treatment of blame (in both its private and overt varieties) would take us far afield, on our view, overt blame is private blame's outward manifestation. Private blame is a manifestation of a disposition to *regard* the wrongdoer in a certain (cf. McKenna 2012). According to a very popular view, the relevant kind of regard involves the manifestation of a "reactive attitude," such as resentment, indignation, or disapprobation (e.g., Fischer and Ravizza 1998, Russell 1995, Wallace 1994, and Strawson 1962). Overt blame, then, is a manifestation of a certain kind of disposition to *treat* the wrongdoer in a certain way, such as the examples given above. We should be clear that on our view, to blame someone is not the same as to judge someone blameworthy. One might judge someone blameworthy without being disposed to blame them or manifest such a disposition (see McKenna 2012). Consequently, on our view, forgiving does not require one to give up a judgment of blameworthiness.

[13] On forgiveness typically ruling out both overt blaming and revenge, see, e.g., Butler (1846), Garrard and McNaughton (2002), Murphy (2003), Griswold (2007), Pettigrove (2012), and Warmke (2011, 2013, 2015). Forgiveness probably also makes many or all forms of *private* moral blame inappropriate as well. But because no one would ever confuse punishment (an essentially overt activity) with private blame (an essentially private activity), we will set aside the matter of private moral blame.

theorists claim. What we shall argue is that in many cases of forgiveness, there are behaviors that are morally compatible with forgiveness that can plausibly be taken to be forms of punishment. Of course, our critic might wish to argue that those behaviors we claim are sometimes morally compatible with forgiveness are not actually forms of punishment, and so cut no ice against their theory. We invite them to make those arguments. We take this to be an opening salvo.

With these initial points in mind, let us turn to the arguments we canvassed above. In what follows, we hope to raise serious doubts about the view that forgiving rules out punishment. While we do not take our arguments to be decisive objections, we think that they identify the sorts of considerations that punishment forbearance theorists of forgiveness must contend with if they hope to vindicate their view.

We first considered the forfeiture argument. There is something appealing to the main idea of this argument, which is that in forgiving a wrongdoer, we forfeit our liberty right to punish her (or at least to do so any further).[14] If we have forfeited this right, then it is no longer morally permissible to punish the wrongdoer (for the wrong for which she was forgiven).

The problem with the forfeiture argument, however, is that it is easy to come up with cases in which it looks like while an agent has forgiven a wrongdoer, she has not forfeited her right to punish her. We can imagine a child who lies to her parent (a clear violation of house rules) and does so in front of her siblings. The mother reprimands the child and the child apologizes and expresses sadness and regret over her poor choice. The mother then says the following: "I love you very much, and it hurts me when you lie to me. I'm glad you apologized, though, and I forgive you. However, you know the punishment for lying in this house, so please go to your room for the next hour. I'll let you know when you can play with your toys again."[15] It seems to us that in such a case, the mother can both forgive and yet retain her right to punish the child. If so, then it is not the case that one necessarily relinquishes one's right to punish, thereby making one's punishment of the wrongdoer morally inappropriate.[16]

Turn now to the no justification argument. The crucial claim of this argument is that to forgive one's wrongdoer entails that (among other things) one has dismissed the fact that she wronged you *as a reason* to punish her. The defender of this argument might concede

---

[14] A liberty right to do X is generally understood as a "cluster right," as it contains both a permission to do X, and a claim right against interference in doing X. See Thomson (1990, pp. 53-6).

[15] Hughes (1993) suggested such a case in a different context and Warmke (2013) gave another version of it.

[16] It might be objected that this case relies on a hierarchical relationship between mother and child, and that we had taken ourselves to be widening our discussion of punishment to ordinary interpersonal cases where such hierarchical relationships need not exist. Even supposing we grant the point, what follows from it? If one can forgive without forfeiting one's right to punish when one has the right, does it matter to the forfeiture argument whether the relationship is a hierarchical one or not? We cannot see how, at least not according to the current construal of the forfeiture argument.

that when we forgive we do not forfeit the *right* to punish. But from this it does not follow that punishment is morally appropriate. For when we forgive, we decide not to act on certain classes of reasons. We take certain reasons for action "off the table," as it were, even if, strictly speaking, we retain a right to act for those reasons.[17] Upon forgiving, then, one lacks justification to punish because one has disavowed the kinds of reasons that would naturally motivate one to punish a wrongdoer. Having those reasons for punishment undercut, punishment becomes inappropriate.

There is something attractive about this argument, too, as it is natural to think of forgiveness as a way of deciding no longer to be swayed by a certain class of reasons. In so doing, it seems plausible that agents who forgive might let go of reasons that would have otherwise motivated them to punish their wrongdoers. Perhaps the desire to "get even" will evaporate entirely, relying as it does on a kind of personal indignation toward one's wrongdoer. Minimally, such desires will be disavowed and not endorsed—not seen as acceptable reasons for action.

We think, however, that the no justification argument depends on a mistakenly narrow view about what can count as justifying reasons to punish a wrongdoer. For it seems that one can let go of one's feelings of anger, hurt, personal protest, bitterness, or rancor toward the wrongdoer, and yet insist upon punishing her out of other moral considerations. For one simple way of seeing this, consider two ways of expressing a moral complaint. One might say to a wrongdoer "you lied to me." This couches the complaint in personal terms, and emphasizes one's indignation at having been wronged by this person—as having done something to *this relationship* (see, e.g., Scanlon 2008). On the other hand, one might say to the same person "you lied." This is an impersonal, or general, expression of a moral complaint (even though the wrong was done to the one making the complaint).[18] The general expression of the complaint seems to place the wrong in a sphere of broader concern: it casts the wrong as a kind of offense against everyone (or perhaps against God, or the moral order, or whatever). The general form of the complaint emphasizes not the directional feature of the wrong (i.e., "to me"), but the fact of the wrongdoer's having committed it. Most importantly, for our purposes, the general complaint seems to be much more in line with forgiveness than does the personal complaint.[19] Thus, you might forgive the wrongdoer for lying *to you*, yet still permissibly persist in regarding or treating her as someone *who lied*.

The deep moral point behind all this is that there are many kinds of moral reasons. Even if we grant that forgiveness nullifies the motivation for making the *personal* form of a moral complaint (that would typically be at issue when we blame *as victims* of wrongdoing), it leaves intact the fact that the wrong occurred and so does not interfere with the grounds of the *general* moral complaint (the kind of complaint that third parties who were not

[17] Cf. Waldron (1981).

[18] We do not claim that no one would express a general moral complaint by uttering dispassionately "you lied to me." But these seem to us typical ways of expressing the two forms of complaint.

[19] Another way of expressing the point might be to say that forgiveness ought not be thought of as a matter of justice, but that punishment should (see Griswold 2007, pp. 32-3).

victims of the lie could *share* with victims). And if the general complaint stands, then it can also serve to motivate and justify punishment of the wrongdoer. And so even if the victim of wrongdoing lets go of her grounds to make personal complaints against the wrongdoer, this would not entail that she has also let go of her grounds of making general or impersonal complaints, complaints that could be shared by anyone for whom indignation can be an appropriate response. These general grounds for complaint could then persist through forgiveness: by letting go of one's personal complaint, then one need not also let go of any and all justifications to punish.[20]

For another type of response to the no justification argument, we can consider the standard justifications of punishment. These are typically advanced as defenses of institutional punishment, of course, but the same reasons behind them may apply to cases of interpersonal punishment (whatever we think of the wisdom of actually attempting to apply them). For each justification, we can ask whether the fact of the victim having forgiven her wrongdoer interferes with the justification for punishment. If the justification for punishment remains, then that is some evidence that, by forgiving, a victim need not be understood as disavowing or taking such reasons "off the table."

Take first consequentialist justifications of punishment. The aim of punishment on these views is to bring about good consequences, and especially to deter further wrongdoing from either the wrongdoer herself, or other prospective wrongdoers. Forgiving a wrongdoer does not seem to remove the need to deter further wrongdoing, by her or others.[21] It might be objected that if the victim forgives the wrongdoer, then there will be no need for further deterrence (at least in the wrongdoer's own case), as perhaps the victim is well-positioned to know whether the wrongdoer needs to be deterred, or has really reformed.

In reply, we note that on the theory of punishment under consideration, either punishing a wrongdoer in order to deter her from wronging again is the optimal action in terms of consequences or it is not. If it is, then we have a justification for punishing her, and whether she has been forgiven is irrelevant to whether punishing her is optimal. On the other hand, if punishing her is not optimal, then whether she should be punished or not does not have anything to do with whether she has been forgiven.[22] The upshot is that on this view of punishment, one's forgiveness need not be tied to forbearing punishment. The matter of punishment is simply tied to the matter of promoting good consequences and deterring bad ones. These reasons can remain for a victim can remain, we claim, even if she forgives.

---

[20] Paul Hughes (1993) appears to have a similar sort of thought in mind: the fact of wrongdoing as such can remain as a grounds for punishment, even when we forgive.

[21] This may be true even if the justification for moral blame is also consequentialist (e.g. Vargas 2013), for the aims of each of these practices may be to secure different kinds of ends.

[22] Whether the wrongdoer has been forgiven is, of course, even less relevant to whether punishing her would be an efficient method for deterring wrongdoers in general.

Consider next the communicative account of punishment.[23] On this family of views, punishment is a communicative act that, roughly, sends a message to the wrongdoer, victim, and broader community that the victim did not deserve to be wronged, and that the wrongdoer is culpable for doing so. It seems to us that on this view, forgiveness makes punishment all the more morally urgent. After all, the victim risks sending the message that it is acceptable to mistreat her if she simply forgives her wrongdoer without some accompanying message to make clear the fact that the wrong was unacceptable.[24] Perhaps it will be objected that coupling forgiveness with an expression of commitment to punishment sends a mixed, or even incoherent, message.

To this we reply that the message need not be mixed or incoherent. Recall the mother-child case from above, in which the mother says: "I love you very much, and it hurts me when you lie to me. I'm glad you apologized, though, and I forgive you. However, you know the punishment for lying in this house, so please go to your room for the next hour. I'll let you know when you can play with your toys again." It seems to us that in such a case, the mother is not sending mixed messages. While she need not deploy something as complicated as the communicative theory of punishment in her explanation to her child, the child can understand that, while she is forgiven for her lie, she is being punished for breaking the house rules (the corresponding punishments for breaking we can imagine being justified by communicative theories).

For retributivists, the purpose of punishment is to give wrongdoers their just deserts.[25] Retributivists therefore need to specify a desert-basis—an answer to the question of what makes a person deserve punishment to a certain degree. The desert basis could plausibly be argued to be any number of things—the degree of culpability of the wrongdoer's choices, the extent of harm she causes, or even the degree of wickedness of the wrongdoer's character. One thing that could not plausibly determine what wrongdoers deserve, however, is the degree of resentment harbored against them by their victims that is discharged upon forgiving. Of course, we might hope that rational people could make reasonable assessments of what their wrongdoers deserve, but then of course the assessments themselves would be at best a proxy for the appropriate moral considerations. And even if the victims were uniquely situated to determine accurately the degree of deserved punishment, it is implausible to hold that a victim's assessment itself *determines* the answer.[26] So since the desert-basis must be something other than the degree of victim resentment (or whatever other kind of attitudes taken up by the victim) that is discharged upon forgiving, the fact that a victim forgives a wrongdoer need not undermine one's justification for punishing that wrongdoer.

---

[23] See, e.g., Duff (2001).

[24] There is, of course, a concern that the message need not be delivered specifically through punishment, but that is a general quibble with the communicative view of punishment, and not one raised specifically by cases of forgiveness.

[25] See, e.g., Moore (1987), Morris (1968), Murphy (1973).

[26] It would be much more plausible to hold that the morally appropriate level of resentment felt by a victim tracks the degree of fitting punishment. But then it is not the victim's resentment that determines the degree of fitting punishment (if any), but whatever makes that resentment morally appropriate.

The restitution model of punishment might seem to lend support to the punishment-forbearance view of forgiveness.[27] The purpose of punishment, on that model, is not for the wrongdoer to suffer unpleasant consequences, but rather for him to restore the victim as best as possible to the condition she was in before he wronged her—in other words, to pay a debt to victim. So it could be that if the victim forgives the wrongdoer, she also absolves the wrongdoer of any debt. And since punishment consists solely in the repayment of the debt arising from a wrong, there is no remaining justification for punishing the wrongdoer.

However, we see two problems with this line of argument. First, even if I, as a victim, forgive your interpersonal moral debt to me, this does not entail that *all* of your debts that you incurred by your wrongdoing have been cancelled. There may be debts remaining that I as the victim do not even have the standing to cancel through my forgiveness. To hold that the debt is owed only to the victim is to construe the nature of wrongdoing rather narrowly, as a purely private affair. It seems more plausible to us that wrongdoers also typically incur a more general debt to the moral community, so that the victim is not the only party with a claim against the wrongdoer. A wrongdoer might pose a similar risk to others, or his action could be understood as a wrong against all those who follow the rules of the moral community.[28] If the victim sometimes has the standing to punish the wrongdoer because of these remaining debts, then even if forgiveness involves the cancellation of certain kinds of personal moral debts, this would not require the victim to forbear punishing the wrongdoer, so long as she has standing to do so. (Of course, if the victim does not have the standing to punish in order to secure payment of these outstanding debts, then the question of the permissibility of punishment after forgiveness has already been settled and the explanation will have nothing to do with having forgiven her.)

The second problem is that, even if we focus on the debts that can be cancelled by the victim of the wrongdoing, we need not think that, by forgiving, the victim necessarily cancels *all* debts the wrongdoer owes to her. Perhaps forgiveness cancels some kinds of interpersonal debts, but not others. You might, in a fit of rage, drop a can of paint on your neighbor's living room carpet. Your neighbor might forgive you for this mean act, but still be, we think, within her rights to ask you to replace her carpet. That debt to her still might remain. The point, then, is that if forgiveness does not necessarily cancel *all* debts incurred because of a wrong, then some of those debts that remain might be ones that can justify punishing the wrongdoer. The defender of the view under consideration must claim, by contrast, that interpersonal forgiveness cancels any and all debts that could plausibly serve as a justification for punishment according to restitution theories of punishment.

Next, consider the moral education account of punishment. On this view, the purpose of punishment is to reform the character of the wrongdoer. Presumably no one would think

[27] The most prominent defender of this view is Barnett (1977).

[28] The idea of a general debt to society is useful, and perhaps indispensable, for explaining crimes without individual victims. Cf. Dagger (1980, 1991)

that forgiveness itself will necessarily cause a wrongdoer's moral reform—though perhaps it could humble some wrongdoers and start them on a path to redemption. And so despite being forgiven, a wrongdoer may still require further reform-based punishment.[29] If so, then there would be reasons to punish that need not be disavowed when one forgives. Someone could object that forgiveness is only appropriate when a wrongdoer has fully reformed. But if forgiveness requires that the wrongdoer already be reformed, then it would be wrong to punish a fully reformed person whether she has been forgiven or not. The justification for her punishment would disappear along with her vicious character, and this disappearance would have nothing to do with first being forgiven.

Finally, take the new relationship argument. This argument casts forgiveness as a way of morally repairing a relationship. If forgiveness is (at least in part) about reconciliation with one's wrongdoer, then it does seem a bit odd to say that you forgive someone who has wronged you, yet continue to punish her all the same. How serious could you be about reconciliation, after all, if you take one step toward repair and then immediately jump backwards? To this argument we answer that people seem to have no trouble doing exactly this in practice, and it seems sometimes to be a great relief to both victim and wrongdoer. It would not be unusual to hear of someone asking for nothing more than forgiveness, even if it is granted that the relationship will never be the same, or that victim and wrongdoer will never see each other or speak again. Yet if we insist that forgiveness requires forbearance from punishment because forgiveness must always involve seeking reconciliation, these requests become difficult to understand, amounting as they do to asking for forgiveness while leaving punishment in place. While forgiveness is typically aimed at full or partial reconciliation, there is no quick easy path from forgiveness to reconciliation, and in some cases, there may be no clear path at all.

Here is a second practical counter-argument. Forgiveness is generally not regarded as an outrageous thing to ask for. It is a part of every twelve-step program. It is often (and hopefully!) asked for with a great deal of humility. Yet if asking for forgiveness necessarily includes asking for forbearance from punishment on the grounds that the victim should not do anything that does not directly promote reconciliation, then it is not hard to see that asking to be forgiven would often be a truly outrageous request.[30] Far fewer people would have the temerity to ask for it, and so they and their victims would miss opportunities for a partial moral repair that many see as something worth wanting. So it turns out that far from being odd, we have a great deal of practical use for forgiveness if it requires only a partial moral repair, and relatively little use for it if it must require avoiding anything that could impede reconciliation, such as subsequent punishment.

A third reply rests on pointing out an implied assumption in the new relationship argument: that punishment cannot be an important means to reconciliation. But why think that? Indeed, being punished may not always feel in the moment as if it is promoting reconciliation—it hurts. But getting a rabies vaccine hurts, too. This does not

[29] Cf. Duff (2001, pp. 107-9).

[30] "Listen, I know I stole your car, slept with your brother, and ran up $100,000 in credit card debt in your name, but my sponsor says I have to ask, so how about we pick up where we left off?"

rule out the vaccine being a necessary means to future health. Punishment, too, might not rule out future reconciliation—indeed, it might be a very important (perhaps even necessary) stopover on the long road to reconciliation. If that is correct, then punishment might itself be a vital part of the new post-wrongdoing relationship inaugurated between victim and wrongdoer.

In this section we have argued that the forfeiture argument, the no justification argument, and the new relationship argument each fail to vindicate the claim that forgiving a wrongdoer morally requires one to forbear punishing her. This is so even if we grant that punishment is something individuals commonly have the standing to do in their interpersonal relationships. If correct, we think this undermines at least some obvious strategies for arguing that forgiveness makes subsequent punishment impermissible. It is worth noting that this result does not directly cut against the claim that forgiving *itself* involves something like the forswearing or forbearing of punishment. To argue against the moral claim that forgiveness morally requires forbearing punishment is not to argue against the claim that to forgive *is* (in part or whole) the forbearance of punishment. Of course, these views are connected: if it is false that forgiveness makes punishment morally impermissible, then that would be very good evidence that it is also false that, as a necessary condition on forgiveness, one must forbear punishment.

4. Institutional Forgiveness?

We noted earlier that punishment is a practice that we generally associate with institutions, and that it is thus odd to think of forgiveness as requiring one to forbear punishment, as if this was typically an option for individuals. We made sense of the view by introducing an expansive interpretation of punishment, which included interpersonal punishment. But perhaps the punishment-forbearance view is best studied in a different context. Suppose we maintain a more limited, institutional view of punishment, but also think of forgiveness as something that institutions can do. In fact, some theorists have done exactly that (Digeser 2001). Perhaps thinking of forgiveness at the institutional level will reveal something about the concept of forgiveness that is not immediately clear when we think about interpersonal forgiveness.

The first institution that comes to mind when one thinks of punishment is the state. If the state were to forgive a wrongdoer, what would that involve? Issuing a pardon, and so foreclosing the possibility of further punishment for the wrong in question, is one obvious possibility. But might the state do less and still be said to have forgiven the wrongdoer? Suppose that some public official responds to a wrongdoer's request for clemency by issuing a statement expressing sympathy for her plight and a warm assessment of her improved character, but nonetheless reaffirms the plan to punish her. One possible assessment of the case is to say that this is not forgiveness at all, as it might be a pure political move made in response to pressure from some interest group. But suppose we specify that the official's order is only made as a result of rethinking the state's relationship to the wrongdoer, in isolation from political-careerist concerns.

Is this a case of forgiveness? Issuing such a statement seems, at least to us, like an odd thing for the official to do. Perhaps we can make progress by asking whether the

statement described is something the wrongdoer would value. If she would value the statement, then that would suggest that her relationship with the institution of the state is something she values apart from any punitive consequences attached to it. And if she does so value the relationship, then that suggests that we have a practical use for an understanding of institutional forgiveness that does not include forbearing punishment. But unless she has either a close relationship with the public official or is civic-minded in a way not commonly seen in modern states, the answer is likely 'no.' In general, we suspect that practically anyone in a position to be forgiven by the state is interested only in punishment forbearance.[31]

Can this feature of state forgiveness be generalized to any other plausible forms of forgiveness? We think not. It cannot be generalized to interpersonal forgiveness, because, as we have already argued, people do typically have a practical interest in forgiveness in interpersonal relationships that does not include punishment forbearance. More plausibly, we might consider the possibility that there are different kinds of forgiveness for different kinds of agents.[32] It might be thought that institutional forgiveness in general (and not just by the state) includes punishment-forbearance. But this also seems false. It may be odd to think that citizens of modern states would value forgiveness without punishment-forbearance, but things may have been different when peoples' identities were more closely bound up with their political community. And states are hardly the only institutions that can engage in punitive behavior. A person might plausibly care about being forgiven by a religious organization, charity, hospital, or any number of other kinds of institutions with which she might have fallen out. So if we want to hold out the possibility of forgiveness from these sorts of institutions, we should note that punishment forbearance does not seem necessary.

We close this section by asking whether state forgiveness is a useful concept at all. Perhaps it is, but it does not seem to add anything to our familiar way of talking about state pardons. In fact, the idea of forgiveness seems rather out of place in the context of talking about a type of institution to which few still have a deep emotional connection akin to a personal relationship. Ironically, it is this lack of emotional depth in the citizen-state relationship that makes the requirement of punishment-forbearance in "state forgiveness" plausible in the first place: without it, there would be very little state action that is even in the neighborhood of forgiveness. So it turns out that there is a lesson to be learned by considering the possibility of institutional forgiveness that is relevant to interpersonal

[31] It might be objected that just because the wrongdoer has no interest in being forgiven, this does not mean that she is not forgiven, or that such phenomena should not count as a kind of forgiveness. After all, even in ordinary, interpersonal cases of forgiveness, sometimes the wrongdoer has no interest in being forgiven, either because she thinks she did nothing wrong, or because she does not value her relationship with the putative victim of her wrongdoing. But the difference between these cases, we think, is that, generally speaking, people *do* care about being forgiven in ordinary interpersonal contexts, whereas generally, people would not usually care about being "forgiven" by the state (or a representative thereof) unless that forgiveness involved pardon or something similar. So while we grant that this "value argument" does not show that states cannot forgive over and above simply forbearing punishment, it does draw a disanalogy between state forgiveness and ordinary forgiveness that should give pause.

[32] Digeser distinguishes between "ordinary forgiveness" and "political forgiveness" (2001, p. 28).

forgiveness, but it is not the kind of lesson for which punishment-forbearance theorists might have hoped. This discussion suggests that punishment-forbearance accounts of forgiveness are most plausible when we remove the emotional complexity of interpersonal relationships. But we worry about the prospects for a claim about forgiveness if the claim is made most plausible in contexts that have very little in common with the interpersonal relationships in which forgiveness is typically thought to find its paradigmatic instantiations.

5. Forgiveness and Mercy

Thus far, we have argued against a conception of forgiveness that requires one to forbear punishment. Of course, we do not mean to argue that if one forgives, one *must* punish one's wrongdoer (assuming such a thing is even possible in non-hierarchical relationships). All we mean to have shown is that forgiveness can be morally consistent with punishment. Naturally, however, one might wish to forbear punishment of one's wrongdoer—and this may or may not be concomitant with one's forgiving the wrongdoer. In other words, forbearing punishment might be an act of mercy, but this act of mercy is separable, conceptually and morally, from one's having forgiven the wrongdoer. We conclude, then, by reflecting briefly on some key differences between forgiveness in particular, and acts of mercy in general.

First, manifestations of mercy, but not forgiveness, are essentially overt. To extend or show mercy to someone who has acted badly is to engage in some overt behavior: a governor may show mercy by commuting a guilty criminal's sentence or a parent may lessen a guilty child's punishment. In any such cases, however, the act of mercy is overtly expressed. Forgiveness, however, is not necessarily overt. It is possible to forgive privately; indeed philosophical discussions of forgiveness have focused predominantly on its private manifestations.

Second, mercy is third-personal in a way that forgiveness is not. Forgiveness is paradigmatically a second-personal enterprise, whereby the only person in a position to forgive S for S's wronging of P is P herself. As some philosophers have suggested, there are perhaps exceptions to this general rule, but even so, these cases are clearly non-paradigmatic.[33] Mercy, on the other hand is very often shown third-personally. A governor need not be the victim of a criminal's offense in order to show mercy and commute a sentence. A parent may abate Little Tommy's punishment even though Tommy's offense was to hit his sister. A boss may reprimand an employee for telling an inappropriate joke in the lunchroom even though termination would be justified and the boss herself is not personally offended by it. In all of these cases, a third-party shows mercy, and does so (we may presume) justifiably. This is not to say that mercy *must* be third-personal. After all, the boss herself might have been personally offended by a joke aimed specifically at her. But forgiveness, if it is ever third-personal, is not third-personal in *this* way. Barring exceptional circumstances, you cannot forgive me for the offense I caused to a co-worker—you simply lack the standing to forgive me for my offense to *her.* Mercy is not paradigmatically second-personal in this way.

[33] See, for example, Griswold (2007, pp. 117-119).

Third, mercy is often (if not always) connected to authority in a way that forgiveness is not. Indeed, our examples above were instances of mercy shown by individuals in positions of legal, parental, and occupational authority. It is natural to think that in order for S to show mercy to P, S must be in *some* kind of position of authority over P. You might even be justifiably offended if someone who is *not* in a position of authority over you was to communicate her desire to show you mercy for something you had done—you might take her comment as an implicit assertion of her authority over you, something you might justifiably resent if she is your equal. Interpersonal forgiveness, however, is tied to no such authority structures: peasants may forgive kings.

Whatever similarities they possess on the surface, forgiveness and mercy are clearly distinguishable phenomena. Mercy is essentially overt, paradigmatically third-personal, and tied to authority. Forgiveness, however, is possibly private, paradigmatically second-personal, and independent of authority.

6. Conclusion

We began by speculating that there is a common folk view that forgiveness and punishment are somehow linked, such that one should not punish a person after forgiving her. We think that the discussion above presents a strong prima facie case for rejecting that view. Forgiveness, acts of mercy, the withdrawal of blame, the forbearance of revenge, and the forbearance of punishment are indeed moral phenomena that reside in the same neighborhood. It is understandable that some might draw strong connections between any and all of them. Indeed, we think that some of them are in fact connected. Though we have not argued for it here, we do think that forgiveness typically morally requires one to forbear revenge (and most, if not all) acts of moral blame. But even if this is so, it does not follow that forgiveness also morally requires one to forbear punishment. Indeed, we have tried to show that even if victims do indeed have the authority to punish, there can be plenty of justifying reasons for a victim to punish her wrongdoer after forgiving her. Further, forgiving itself need not involve the forfeiture of a liberty right to punish. And even further, the fact that one has forgiven does not entail that subsequent punishment violates the normative standards of a post-forgiveness relationship. Of course, it goes without saying that we are not claiming that just *any* kind or measure of punishment can be appropriate after forgiveness. It seems to us, however, that the moral relationship that punishment-forbearance theorists claim holds between forgiveness and punishment is not as strong as advertised.[34]

Bibliography

Barnett, Randy E. 1977. "Restitution: A New Paradigm of Criminal Justice." *Ethics*, 87(4): 279-301.

[34] We thank Jonathan Jacobs for his excellent comments on a previous draft of this chapter.

Bash, Anthony. 2015. *Forgiveness: A Theology.* Cascade Books: Eugene.

Boonin, David. 2008. *The Problem of Punishment*. New York: Cambridge University Press.

Butler, Joseph, 1846, *Fifteen Sermons Preached at Rolls Chapel, in The Works of the Right Reverend Father in God, Joseph Butler, D.C.L., Late Bishop of Durham,* Samuel Halifax (ed.), New York: Carter.

Dagger, Richard. 1980. "Restitution, Punishment, and Debts to Society." in *Victims, Offenders, and Alternative Sanctions*. Joe Hudson and Burt Galaway (eds.). Lexington, Mass.: Lexington Books: 3-14.

_____. 1991. "Restitution: Pure or Punitive?" *Criminal Justice Ethics*. 10(2): 29-39.

Digeser, P.E. 2001. *Political Forgiveness*. Ithaca: Cornell University Press.

Duff, R. A. 1986. *Trials and Punishments*. Cambridge: Cambridge University Press.

_____. 2001. *Punishment, Communication, and Community*. Oxford University Press.

Enright, R.D., Eastin, D.L., Golden, S., Sarinopoulos, I., and Freedman S. 1992. "Interpersonal Forgiveness Within the Helping Professions: An Attempt to Resolve Differences of Opinion," *Counseling and Values* 36: 84-103.

Fischer, John Martin and Martin Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Garrard, Eve, and David McNaughton, 2002, "In Defence of Unconditional Forgiveness," *Proceedings of the Aristotelian Society*, 103 (1): 39–60.

Gert, Bernard. 2010. *Hobbes: Prince of Peace*. Cambridge: Polity.

Griswold, Charles L. 2007. *Forgiveness: A Philosophical Exploration*. New York: Cambridge University Press.

Haber, Joram Graf, 1991, *Forgiveness*. Lanham, MD: Rowman and Littlefield.

Hart, Herbert. 1968. *Punishment and Responsibility*, Oxford: Oxford University Press.

Hobbes, Thomas. 1651. 1969 edition, *Leviathan*. Menston, England: The Scholar Press Limited.

Hughes, Paul M. 1993. "What is Involved in Forgiving?" *Journal of Value Inquiry*, 27: 331–340.

Jacobs, Jonathan. Forthcoming. "Resentment, Punitiveness, and Forgiveness: An Exploration of the Moral Psychology of Punishment," *This Volume.*

Londey, David. 1986. "Can God Forgive Us Our Trespasses?" *Sophia* 25: 4-10.

Mabbott, J.D. 1939. "Punishment," *Mind*, 48(190): 152-167.

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

Morris, Herbert. 1968. "Persons and Punishment." *The Monist*, 52: 475–501.

Murphy, Jeffrie G. 1973. "Marxism and Retribution." *Philosophy and Public Affairs*, 2(3): 217–43.

_____. 2003. *Getting Even: Forgiveness and its Limits*, New York: Oxford University Press.

Moore, Michael S. 1987. "The Moral Worth of Retribution." in *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, ed. Ferdinand David Schoeman. 179-219. Cambridge: Cambridge University Press.

O'Shaughnessy, R.J. 1967. "Forgiveness," *Philosophy*. 42(162); pp. 336-352.

Pettigrove, Glen. 2012. *Forgiveness and Love.* Oxford: Oxford University Press.

Russell, Paul. 1995. *Freedom and Moral Sentiment. Hume's Way of Naturalizing Responsibility.* Oxford: Oxford University Press.

Scanlon, T.M. 2008. *Moral Dimensions: Permissibility, Meaning, and Blame*. Cambridge: Belknap Press.

Simmons, A. John. 1991. "Locke and the Right to Punish." *Philosophy and Public Affairs*, 20(4): 311-349.

Strawson, P.F. 1962. "Freedom and Resentment" in *Free Will: Second Edition*, Gary Watson (ed.). 2003. Oxford: Oxford University Press.

Swinburne, Richard. 1989. *Responsibility and Atonement*, New York: Oxford University Press.

Thomson, Judith Jarvis. 1990. *The Realm of Rights*. Cambridge, Mass.: Harvard University Press.

Vargas, Manuel. 2013. *Building Better Beings: A Theory of Responsibility*. New York. Oxford University Press.

Waldron, Jeremy. 1981. "A Right to Do Wrong." *Ethics*, 92(1): 21-39.

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

Warmke, Brandon. 2011. "Is Forgiveness the Deliberate Refusal to Punish?" *Journal of Moral Philosophy* 8(4): pp. 613-620.

_____. 2013. "Two Arguments Against the Punishment-Forbearance Account of Forgiveness," *Philosophical Studies* 165 (3): pp. 915-920.

_____. 2015. "The Economic Model of Forgiveness," *Pacific Philosophical Quarterly*.

Wellman, Christopher Heath. 2009. "Rights and State Punishment." *Journal of Philosophy* 106(8): 419-439.

Wolterstorff, Nicholas. 2009. "Jesus and Forgiveness," *Jesus and Philosophy: New Essays*, Paul Moser (ed.). Cambridge: Cambridge University Press.

Zaibert, Leo. 2009. "The Paradox of Forgiveness," *Journal of Moral Philosophy* 6(3): pp. 365-393.