

# Forgiving, Committing, and *Un*-forgiving

Monique Wonderly

University of California, San Diego

**Correspondence**

Monique Wonderly, University of California, San Diego.  
Email: mlwonderly@ucsd.edu

**Abstract**

Theorists often conceive of forgiveness as “wiping the slate clean” or something of the sort with respect to the offender’s moral infraction. This raises a puzzle concerning how (or whether) the relevant wrongdoing can continue to play a role in the forgiver’s deliberations, attitudes, and practical orientation toward the offender once forgiveness has taken place. For example, consider an agent who forgives her offender for an act of wrongdoing only to later blame her again for that very same act. Is the relevant agent morally criticizable for breaching a commitment internal to her earlier forgiveness? When, if ever, is “un-forgiving,” justified? In brief, I argue that sometimes, we can genuinely forgive a wrongdoer for a particular transgression and at some later time, justifiably un-forgive her for it. Though the act of un-forgiving is underexplored, I argue that attending to this phenomenon will not only illuminate an important, often overlooked aspect of moral life, but it will also help to inform extant accounts of the nature and ethics of forgiveness.

Sometimes, we forgive those who wrong us. In ideal cases, the offender apologizes and expresses remorse, and we, in turn, might “wipe the slate clean” with regard to the wrongdoing, paving the path to move forward. This is one way the story might go, but there are others. As many of us who have had occasion to seek (or to extend) forgiveness know, clean slates can be hard to come by, in part, because acts of forgiveness – and the relationships that they implicate – can be quite messy. Consider the following scenario:

*Victor and Wanda.* Victor and Wanda are in a committed, monogamous relationship, and Wanda has a brief affair with an acquaintance. Victor is deeply hurt and resents Wanda for her transgression. Wanda expresses regret, apologizes, and engages in ardent attempts to make amends. She asks Victor to forgive her, and after a time, he finally tells her “I forgive you.” However, several months later, he begins making cutting remarks about

Wanda's act of betrayal. "I can't believe you did that to me!" "Is that the same restaurant you went to with *him*?" "You're late! Whom were you with *this time*?" Wanda protests, "But you said you forgave me!"

Most of us would tend to sympathize with Wanda's frustration. There does seem to be something illicit about Victor's attitudes and behaviors toward Wanda, given his earlier purported forgiveness. On one common view, in forgiving Wanda, Victor committed to refrain from blaming her for the relevant wrongdoing in the future, and it thus appears that he has culpably breached that commitment. As I will argue, however, we should make room for the possibility that though Victor initially forgave Wanda, he later justifiably *un-forgave* her for her offense.

Though the topic has been largely neglected in the philosophical literature, interrogating the justifiability of un-forgiving can help to illuminate the structure and normativity of forgiveness. One plausible understanding of justified un-forgiving stands in tension with certain conceptions of forgiveness. In what follows, I articulate and address this tension. In section 1, I employ extant accounts of forgiveness to construct a rough, working sketch of what forgiving involves. In section 2, I introduce the phenomenon of un-forgiving, show how we might un-forgive for reasons, and argue that un-forgiving is sometimes justified. In section 3, I elucidate a tension between the justifiability of un-forgiving and certain views on which forgiveness involves a strong commitment to refrain from future blame. In section 4, I outline an approach to understanding the normative significance of un-forgiveness that favors a more permissive view of un-forgiving. Finally, in section 5, I offer concluding remarks.

## 1 | FORGIVENESS

In order to have a fruitful discussion about un-forgiving, we need to start with a rough, working conception of what it means to forgive. This task requires that we venture into the murky waters of the relevant philosophical literature, but for the sake of brevity, I will remain, more or less, at the shallow end.

Contemporary philosophical accounts of forgiveness tend to draw heavily on a view that is often attributed to Joseph Butler – namely, that to forgive is to forswear or overcome resentment.<sup>1</sup> Theorists have elaborated on this notion of forgiveness in various ways. Some expand the *scope* of negative attitudes that forgiveness overcomes to include contempt, dejection, grief, or even "all hard feelings."<sup>2</sup> Others describe the relevant change in attitude as overcoming a broader negative point of view about the wrongdoer – one in which the wrongdoer is seen as "indecent" or otherwise "lowered" in some way on account of her act.<sup>3</sup> Still others focus on the *conditions* or *reasons* that render forgiveness ap-

<sup>1</sup>Theorists generally cite Butler's (1897) Sermon IX, "Upon Forgiveness Injuries," as their source. As Ernesto Garcia (2011) points out, however, this is a misattribution – Butler's actual claim is that forgiveness involves overcoming excessive resentment, in which case resenting would be consistent with forgiving.

<sup>2</sup>Norvin Richards suggests that forgiveness involves overcoming all "hard feelings" (1988, 80). Macalester Bell (2008), along with Eve Garrard and David McNaughton (2011, esp. 97-98), include contempt on their lists. Jeffrey Blustein argues that forgiveness sometimes must overcome emotions such as dejection and grief (2014, 34).

<sup>3</sup>Jean Hampton (Murphy & Hampton, 1988, 38) and Lucy Allais (2008, 50) advocate such a view. For a related account, see Cheshire Calhoun's (1992) "Changing One's Heart." We should also note that some theorists emphasize the forgiver's change in behavior, rather than merely her change in attitude. For example, Brandon Warmke and Michael McKenna write, "One can...engage in behavior that communicates to the forgiven that one will no longer hold the forgiven's blameworthy act against her. And this can be done without the forgiver experiencing antecedent resentment or moderating or eliminating it" (2013, 203.) Similarly, according to Jeffrey Blustein, forgiveness, like blame, can consist largely in behavioral changes (2014, 45).

appropriate. Some, for example, hold that the relevant negative attitude(s) must be overcome “on moral grounds,” or can be overcome for reasons that intelligibly undermine an implicit judgment within the attitude(s).<sup>4</sup> Some add that forgiveness requires not only forswearing certain negative attitudes, but replacing them with positive ones, such as “benevolence,” “good will,” “trust,” or some other “pro-attitude” toward the wrongdoer.<sup>5</sup> Some theorists emphasize certain respects in which forgiveness changes the position or status of the offender. Forgiveness involves “cancelling the offender’s debt to the victim,” “re-accepting” her into a relationship, “restoring moral equality” between the offender and the victim, or otherwise altering the offender’s “normative situation.”<sup>6</sup>

Notice that though we might think of these features as capturing a single, unified notion of “genuine” forgiveness, we might instead think of them as representing *varieties* of forgiveness, where different clusters of qualities attach to distinct phenomena (e.g., affective forgiveness versus relationship-restoring forgiveness). I suspect that our promiscuous use of the “forgiveness” label makes it clear that in practice, we often mean different things when we deploy the term. For example, we can make sense of both the person’s sincere claim that she has forgiven – and yet, is leaving – her abusive spouse and her spouse’s confused plea, “I thought you forgave me?!” as she walks out the door. Sometimes, the kind of forgiveness that we seek differs from the kind that we receive, even where both types are plausibly called forgiveness.

For my part, I make no attempt to offer a theory of forgiveness or to impose constraints on how we use the term. Rather, I employ the views above to help carve out one rough and incomplete, though hopefully familiar, conception of forgiveness. Notice that what is common to most of the preceding views is that they construe forgiveness, in part, as overcoming – or at least mitigating – an adversarial stance that one inhabits in response to the offender’s wrongdoing.<sup>7</sup> So we can begin with the notion that forgiving, whatever else it is, involves the victim ceasing to stand against the offender for the relevant wrongdoing in the same manner that she did prior to forgiveness. To this, we can add that forgiveness is typically viewed as path-dependent, such that the relevant change comes about in some way that distinguishes it from other phenomena like forgetting or mere “letting go.” This vague, broad level of description affords us a useful skeletal framework for understanding forgiveness, but we still yet can put more flesh on these bones.

---

<sup>4</sup>Jeffrie Murphy (Murphy & Hampton, 1988, 24) and Charles Griswold (2007, 40) suggest that forgiveness must be done on moral grounds. Pamela Hieronymi argues that forgiveness responds to reasons to overturn one’s judgment (implicit in resentment) that the offender’s act represents a threatening claim (2001, 552). Per-Erik Milam argues that forgiveness is done in response to a perceived change of heart by the offender (2019, 242).

<sup>5</sup>Robert C. Roberts suggests that forgiveness requires benevolence (1995, 293). Eve Garrard and David McNaughton argue that forgiveness must involve good will toward the wrongdoer (2011, 97-98). Adrienne Martin suggests that forgiveness involves regaining trust (2010, 550). On Glen Pettigrove’s account, forgiving includes a commitment to the well-being of the wrongdoer (2004, 379). Jeffrey Blustein (2014, 95) and Luke Russell (2016, 4) suggest that the forgiver must adopt some unspecified “pro-attitude” toward the wrongdoer.

<sup>6</sup>For discussions of forgiveness as releasing another from a debt, see P. Twambley (1976, 79), Hampton (Murphy & Hampton, 1988, 38), Christopher Bennett (2003), Dana Nelkin (2013, 175), and Brandon Warmke (2016a). Aurel Kolnai (1974, 104), Bennett (2003, 133), and Martin (2010, 551) suggest that forgiveness involves re-accepting the offender into a relationship (of some kind). Hampton (Murphy & Hampton, 1988, 38) and Luc Bovens (2008, 234; 2009, 230) emphasize the role forgiveness plays in restoring equality between the victim and wrongdoer. David Owens (2012, 53), Warmke (2016b, 689), and Bennett (2018, 10) all argue that (at least one form of) forgiveness alters the space of normative burdens (obligations, aptness of blame, etc.) between the victim and the offender.

<sup>7</sup>Note that Luke Russell invokes the notion of an adversarial stance in his interesting argument that forgiveness is sometimes compatible with continuing to punish. He writes, “The victim adopts an adversarial stance towards the perpetrator when she treats the perpetrator in a negative way that the perpetrator would be disposed to challenge, were he free to do so” (2016, 8).

If there is more than one species of forgiveness, in what follows, I focus on the sort for which a shift in affective attitude is central to, if not necessarily exhaustive of, forgiving.

Following the literature, I will sometimes loosely speak of forgiving as, *inter alia*, ceasing to inhabit the adversarial stance constitutive of resentment, the reactive attitude paradigmatically experienced by the victim toward an offender who has wronged her. Furthermore, the relevant brand of forgiveness is sensitive to reasons, and the reasons for which one ceases to resent the wrongdoer will serve to distinguish it from related phenomena. For example, ceasing to resent because my wrongdoer expresses remorse has a different moral character than, say, ceasing to resent merely as a result of memory loss, or again, because I think it would be therapeutic for me to do so. In only the first case is my attitude shift rooted in a reason that implicates the relation in which the wrongdoer stands to the offense and directly grounds the fittingness of my new orientation toward that relation (and its individual relata).

These preliminary remarks are not intended as a complete account of forgiveness, but rather as a starting point for an inquiry into un-forgiving. According to the preceding sketch, the type of forgiveness at issue represents a particular way of resolving *emotional blame*. Un-forgiving, then, will involve a kind of undoing of that resolution. For some, even this thin notion of forgiveness will assume too much. One might think that the causal path that marks out forgiveness has little to do with the victim's reasons, or again, that the victim's emotions are largely irrelevant.<sup>8</sup> For those who hold such views, I invite them to consider the following arguments, insofar as possible, only as they apply to the earlier broad description of forgiveness – or better yet, with one's preferred details filled in. I suspect that any substantive view of forgiveness would benefit from a serious exploration of the possibility of justified un-forgiving.

## 2 | UN-FORGIVING

With a working sketch of forgiveness in hand, we are now better positioned to consider what it might mean to un-forgive. If forgiveness involves ceasing to inhabit an adversarial stance toward the offender with respect to her wrongdoing, then presumably, to un-forgive is, among other things, to re-inhabit that (or a sufficiently similar) stance.

The notion of un-forgiveness may sound strange. The term is uncommon both in philosophical contexts and everyday discourse.<sup>9</sup> Nevertheless, one can find literary examples that at least gesture at the phenomenon at which I am pointing. For example, a biblical passage in the Book of Matthew tells of a king who forgave a servant his debt but “took back” his forgiveness, once he learned that the servant later refused to forgive a third party who was indebted to him.<sup>10</sup> In Robert Hellenga's short story, “The Truth about Death,” the protagonist tells of a curmudgeonly character named Bart, who “never forgave a thing, or if he had, then later he'd take back his forgiveness” (2016, 8-9). And then, there is the 2001 single, “Unforgiven,” in which the American rock band, The Go-Go's, sing the

<sup>8</sup>Interestingly, Dana Nelkin suggests that we can forgive for no reason at all (2013, 170). For views on which affects play a more minimal role in forgiveness, see Cornell (2017) and Bennett (2018).

<sup>9</sup>In the philosophical literature, I have come across only one sustained analysis of un-forgiveness, Geoffrey Scarre's (2016) “Taking Back Forgiveness,” in which he defends the thesis that once one has forgiven, one may not normally take it back. In *Shaping the Normative Landscape*, David Owens remarks that one cannot unforgive (2012, 53). Brandon Warmke briefly considers what he dubs the “retraction objection” as a new challenge to economic models of forgiveness, ultimately leaving the matter unresolved (2014, 583-585). In a forthcoming work, I defend the possibility of un-forgiving (forthcoming).

<sup>10</sup>King James version, Matthew 18: 31-34.

following lyrics: “You’re unforgiven so go on living, knowing that I’ve unforgiven you. And my Thanksgiving came the day I saw it was okay to unforgive you.”<sup>11</sup> Alas, however, Go-Go’s band member and song-writer, Charlotte Caffey, never explained what she unforgave or why it was “okay” to do it. So there still lies some work ahead.

To un-forgive isn’t merely “to blame again” which might occur when a victim forgets and then later remembers a wrongdoing.<sup>12</sup> Consider, for example, the phenomenological differences between being blamed again for a previously forgotten (but not forgiven) act and having one’s forgiveness revoked. In the latter case, the offender is likely to feel not only surprised, but confused and perhaps even slighted. The fact that the victim re-inhabits the relevant adversarial stance after having forgiven the offender matters. Forgiveness marks a positive change in orientation toward the offender and grounds expectations for future attitudes and behaviors. The un-forgiver, then, should be prepared to defend her reversion back to the less agreeable orientation and may be answerable to the offender for it in ways that she wouldn’t have been absent her earlier forgiveness.

Un-forgiving and forgiving are both practices that respond to morally significant features of the world, and both are subject to rational justification. Re-visiting and elaborating on the case of *Victor and Wanda* should help to clarify this point. Imagine, now, that months after Victor forgave Wanda, he heard her bragging to a friend about her first affair and shamelessly plotting another. Her hurtful words “took Victor back” to the initial affair and resentment swelled within him as vivid, painful memories of her betrayal returned. He now finds himself re-inhabiting the same angry feelings and blameful dispositions toward her for the affair that he had when he first learned of it. Cast in this light, it seems to me that Victor justifiably un-forgave Wanda for her initial affair.<sup>13</sup>

To say that Victor was justified in un-forgiving Wanda is to say that he un-forgave her with good reason. I take it that Wanda’s endorsive attitude towards her earlier wrongdoing and her intent to re-offend constitute good reason for Victor to un-forgive her. This idea fits comfortably with a widely accepted view on which the offender’s remorse and a commitment to reform count as reasons that justify, without necessarily rationally requiring, forgiveness.

There are at least two senses in which we might forgive – and later un-forgive – for reasons. If one thinks that forgiveness involves the formation of an intention, or a decision, to forgive, then there is a sense in which forgiving for a reason is similar to voting to impeach a political leader for a reason. In such a case, one performs an action on the basis of some consideration. This sense of forgiving for a reason seems particularly well-suited to some conceptions of forgiveness. For example, if forgiveness is construed merely as a speech act in which the doing of the deed is contained in the utterance, “I forgive you,” then surely we can decide to do it for reasons and directly carry out the act on that basis. On the view I favor, however, forgiveness centrally concerns a shift in affective attitude over which we lack the requisite control to simply enact at will. We can at most decide for reasons *to try* to forgive – e.g., take steps to facilitate the relevant attitudinal change. On a parallel approach to understanding un-forgiving, we might say that Victor deemed Wanda’s current offensive attitude and behaviors as reasons to un-forgive, and so he decided (perhaps merely to try) to un-forgive her on account of them.

---

<sup>11</sup>The song, “Unforgiven” (Armstrong et al., 2001), was co-written with Billie Joe Armstrong and Jane Wiedlin and released from the 2001 album *God Bless the Go-Go’s* by Beyond Records.

<sup>12</sup>Thanks to Craig Agule for helpful discussion on this point.

<sup>13</sup>One might object to my characterization of this case as one in which the victim “un-forgives” his offender. Here, I simply assume that Victor’s change represents the phenomenon I wish to capture and focus on its justifiability. For a defense of the intelligibility of un-forgiving in which I consider alternate interpretations of such cases, including the view that cases of “un-forgiveness” are better construed as cases of conditional forgiveness in which its conditions are violated, see Wonderly (forthcoming).

Some of our attitudes are reasons-responsive, not in the sense that we “decide” to have them for reasons, but in the sense that they are undergirded by reasons-sensitive judgments about, or construals of, the world. To see better how forgiving (and un-forgiving) might be rationally justified in this sense, it will be helpful to consider Pamela Hieronymi’s explanation of how a wrongdoer’s apology might give the victim reason to forgive her. On her account, the wrongdoer’s transgression makes a threatening claim about the victim – namely, that it is “acceptable” to mistreat her (2001, 546). It is this claim that the victim protests via her resentment. In apologizing, the offender renounces the claim, thereby subverting the victim’s judgment that it persists as threatening. Thus, the apology undermines the victim’s resentment and gives her reason to forgive (2001, 548-549). Suppose, though, that the wrongdoer later does something to affirm the claim that he “authored” via his initial act. It makes sense to think that this might give the victim reason to un-forgive. Importing Hieronymi’s framework, then, we might say that in virtue of her recent actions, Wanda renewed Victor’s judgment that her initial transgression makes a threatening claim, rendering the return of his resentment reasonable.<sup>14</sup>

Of course, one needn’t subscribe to any particular theoretical view in order to see how un-forgiving might be justified. One might think that if one forgives while operating under certain false assumptions about – or again, an incomplete understanding of – the relevant wrongdoing, then un-forgiving will likely be warranted. Perhaps I forgave you assuming that you were sincerely remorseful, that you were committed to reform, and that the significance or *meaning* of your wrongful act was, and would remain, within the scope of what I think I can reasonably and responsibly forgive.<sup>15</sup> On the view under consideration, if I learn that these assumptions were mistaken, I might rightly deem my forgiveness an error and justifiably un-forgive you as a means to reversing it.<sup>16</sup>

On my view, regarding one’s forgiveness as a mistake is neither sufficient nor necessary for justifiably un-forgiving. Merely thinking that she forgave sooner than she ought to have done, for example, wouldn’t obviously render the victim justified in un-forgiving her offender. And importantly, I don’t think it necessary that the victim regard her earlier forgiveness as an error in order for her to justifiably un-forgive. To see this, suppose that despite her later change in heart, Wanda was initially remorseful and intended to remain faithful to Victor. Victor might think that his earlier forgiveness was fitting, that Wanda deserved the benefit of the doubt, and that forgiving her was all-things-considered “the right thing to do” to properly respect their relationship. He might, in other words, endorse his earlier forgiveness, even as he now un-forgives Wanda.<sup>17</sup> We can readily see phenomena of this sort at work in cases of what we might call forgiveness’s sister attitudes, love and trust. We sometimes love and trust another with certain expectations (hopes?) in mind where those expectations are disappointed, and we nonetheless think our initial love and trust justified and perhaps very valuable – even where we also think that later withdrawing those attitudes is equally justified.

<sup>14</sup>While I think Hieronymi’s view affords us a useful theoretical framework for explaining the justifiability of un-forgiving, I make no presumption that she would endorse the particular arguments that I present here.

<sup>15</sup>I am indebted to Karen Jones for helpful discussion on this point.

<sup>16</sup>One might deny that “forgiveness” granted under false presumptions counts as actually forgiving (and so does not admit of un-forgiving). See Wonderly (forthcoming) for a discussion and critique of this view.

<sup>17</sup>On this interpretation of Victor and Wanda, Victor continues to believe (correctly) that Wanda *was* sincerely remorseful and did intend to remain faithful. He does not take himself to have been tricked, deceived, or misled by any false information when he initially forgave. One might suggest that his earlier forgiveness was still mistaken insofar as it was based on the false belief that Wanda would remain remorseful and faithful. Yet, it isn’t clear to me that Victor must have held this belief at the time (perhaps he didn’t give much thought to what would happen in the future), let alone that he must have forgiven her because of it. Thanks to an anonymous reviewer from *Philosophy and Phenomenological Research* for prompting me to say more about this here.

One might think that un-forgiving is always justified. On this view, the victim needs no new reasons to un-forgive, as the original offense will always warrant returning to the relevant adversarial stance.<sup>18</sup> In most cases, the victim is never required to forgive, and since she would have needed no further justification to continue resenting the offender, one might doubt that her renewed resentment requires any such justification now. I want to leave it open whether certain acts might be so malicious, brutal, or devastating that the victim can justifiably un-forgive her offender for them without any new reason, but I suspect that such cases will be the exception rather than the rule. Appealing to the initial wrongdoing as sufficient grounds for un-forgiving often will not be enough, as one's forgiveness typically attenuates the reason-giving force of the wrong and counts in favor of not resenting one's offender for it in the future. While a torture victim (for example) might need no new reason to un-forgive her offenders for their atrocity, Victor's un-forgiving does seem to stand in need of justification. Fortunately, as I have argued, he has the requisite justification.

But we need to say more here. If, as some have urged, to forgive is to *commit* oneself to refraining from future blame, then the nature of this commitment will play important roles in determining whether and when un-forgiving is justified. Those who endorse views of forgiveness as involving a strong, directed commitment to the wrongdoer will be able to countenance at best a restrictive view of justified un-forgiving. And those inclined toward a more permissive view of justified un-forgiving might reject accounts on which the commitments internal to forgiveness appear too onerous. The next task, then, is to investigate the type of commitment, if any, that we undertake in forgiving those who wrong us.

### 3 | COMMITTING

An old adage instructs that we *forget* about those offenses for which we forgive. This is one way in which we might "wipe the slate clean," so to speak, with regard to the wronging. While modern theorists rarely counsel, or even countenance, literal forgetting as a part of forgiving, the ideas that forgiveness involves "wiping the slate clean" with respect to the relevant wrongdoing, "putting it behind us," or "letting the matter go," remain very much alive in philosophical discourse. Consider how theorists put the point. Jean Hampton claims that in forgiveness, "the victim 'sends away' the immoral action in the way that a creditor would absolve a debt" (Murphy & Hampton, 1988, 38). Christopher Bennett writes, "In forgiving, a person puts the wrong behind them in some way, determining to leave it out of consideration in their attitude towards the wrongdoer" (2003, 127). On Lucy Allais's account, when you forgive an offender, you "wipe the slate clean" in the sense that "the [wrongdoer's] act is disregarded in your ways of regarding and esteeming her ..." (2008, 56). Notice that these metaphors are not mere descriptions of forgiveness, but they represent ways in which we are *supposed to* "treat" wrongdoings for which we have forgiven. In other words, we tend to think that forgiveness in some sense commits one to forbearing certain blaming behaviors and attitudes in the future.<sup>19</sup>

The idea that we can justifiably un-forgive places constraints on how we interpret the scope and strength of the aforementioned commitments. For example, once the victim ceases to inhabit the

---

<sup>18</sup>Many thanks to Elizabeth Harman for helpful discussion on this issue.

<sup>19</sup>According to Cheshire Calhoun, (expressed) forgiveness conveys to the offender that she "can expect that there will not be later recriminations, cold shoulders, and paybacks" (1992, 77). Luc Bovens suggests that forgiveness involves treating and respecting the offender as a moral equal, adding that once one has forgiven, "resentment-driven agency is uncalled for" (2009, 231). Glen Pettigrove describes the "forgiveness we seek" as one in which the victim gives "reassurance" that "hostile reactive attitudes based on this wrongdoing will remain absent in the future" (2004, 387).

relevant adversarial stance, the cessation needn't endure without end for it to be such that she genuinely forgave. Forgiveness doesn't require that one "wipe the slate clean" in the sense of rendering the forgiven wrongdoing *permanently* ineligible for playing future roles in our attitudes and practices – even those that would constitute blame. This truth is laid bare when we accept the possibility of justified un-forgiving. Relatedly, we can detect a tension between un-forgiving and certain views that interpret forgiveness as involving (what I will call) a strong commitment to refrain from future blame for the relevant offense. In the remainder of this section, I will consider a cluster of accounts of forgiveness that have this character. Each of the relevant accounts represents forgiveness as the voluntary exercise of a "normative power," whereby the forgiver alters the space of obligations between herself and her offender.

Let us begin with the debt-cancellation model of forgiveness, on which forgiving is construed as cancelling the offender's debt.<sup>20</sup> On this view, the forgiven offender is no longer obligated to compensate the victim in certain ways – e.g., via apology and amends, and consequently, the victim is no longer entitled to demand such compensation. As Brandon Warmke helpfully suggests, just as the creditor of a cancelled debt no longer has the right to demand payment from her debtor, the victim who forgives "relinquishes certain rights" to blame her offender (2016a, 585; 2016b, 698). Debt-cancellation models, while perhaps not wholly incompatible with un-forgiving, imply a rather restrictive view of its justifiability.<sup>21</sup> We tend to think of debt cancellations as reversible only in exceptional circumstances. In general, we limit the legitimacy of reinstating previously "cancelled" debt to those cases in which debtor fraud or creditor error render the original "cancellation" non-binding.<sup>22</sup> Notice that such a model could accommodate neither cases in which the gravity of the offense alone suffices to justify un-forgiving, nor cases in which the victim justifiably un-forgives while endorsing her earlier forgiveness. If such cases are plausible, then there seems to be a tension between justified un-forgiving and the commitments involved in debt-cancellation models of forgiveness.

Whereas debt-cancellation models represent the forgiver as, in some sense, committing herself to refraining from future blame (insofar as she voluntarily relinquishes a right to blame), they needn't construe the forgiver as making any explicit commitment to her offender or otherwise.<sup>23</sup> Some views of forgiveness, however, foreground the victim's commitment to treating and regarding the offender in

<sup>20</sup>See Twambley (1976, 79), Hampton (Murphy & Hampton, 1988, 38), Bennett (2003), Nelkin (2013, 175), and Warmke (2016a).

<sup>21</sup>Warmke acknowledges a tension between un-forgiving and debt-cancellation models of forgiveness when he considers "the retraction objection" against such accounts. On his view, one potential route to handling this objection is to suggest that the norms for retracting forgiveness and re-instating a debt that one previously annulled are not as disanalogous as one might suppose. In general, neither should be done, but each admits of exceptions. To illustrate how one might justifiably go back on one's debt-cancellation, he writes the following, "For example, you might have forgiven me a \$100 debt because you believed that my family would suffer greatly if I paid you this debt. Suppose a week later you discover that my family was not really in such dire straits. Your debt-cancellation was premised on a false belief. This might give you sufficient grounds to go back on your cancellation of my debt" (2016a, 15). As I see it, if the debtor was responsible for the creditor's false belief, then this might void the "cancellation." But as I suggest above (and reiterate below), there is reason to think that we can justifiably un-forgive without thinking our original forgiveness an error, and perhaps in exceptional cases, for no new reason at all.

<sup>22</sup>See, for example, Pettigrove (2004, 380, 391 note 48).

<sup>23</sup>Importantly, on Warmke's account, the forgiver's commitment seems to be (at best) secondary, as any such commitment would be unable to explain how forgiveness alters the wrongdoer's obligations (2016a, 588, note 14).

certain ways. I have in mind views on which to forgive is to do something akin to promising.<sup>24</sup> On Geoffrey Scarre's view, for example, the commitment internal to forgiveness is very strong, "amounting to a promise to let bygones be bygones" that the offender should be able to "trust" (2016, 936). On his account, forgiveness can be withdrawn very rarely, and indeed only in the special case where it has been granted in error.<sup>25</sup> Similarly, though Christopher Bennett does not identify the relevant commitment as a promise, on his view, forgiveness does involve a binding, directed commitment to the offender to no longer treat her as standing under certain obligations (2018, 218-219). He labels this aspect of his view, the "commitment thesis," according to which in forgiving, the victim exercises a normative power through which she obligates herself to her offender (2018, 208). Not only does Bennett's view depict un-forgiving as a morally risky affair that is difficult to reconcile with endorsing one's earlier forgiveness, but it ties the permissibility of un-forgiving to the offender's authority and the victim's obligation to *her*.<sup>26</sup> Presumably, the offender can release the victim from her commitment, but if she does not, the victim who un-forgives risks betraying and thereby wronging her offender. In placing the justifiability of un-forgiving on a par with (something like) the justifiability of promise-breaking, Bennett's view, like Scarre's, would seem to set a high bar for justified un-forgiving. I will refer to Scarre's and Bennett's respective views as promissory models of forgiveness.<sup>27</sup>

I have presented the debt-cancellation and promissory models of forgiveness as (what I will call) strong commitment views – i.e., views that construe forgiveness as making a binding commitment to refrain from future blame, on pain of wronging one's offender. But one might deny that the commitment implicit in these views is quite so strong after all and perhaps not even strong enough to seriously strain compatibility with un-forgiving. One might think, for example, that it needn't take much to warrant reversing a debt cancellation or to constitute an exit condition for a promissory commitment.

<sup>24</sup>Glen Pettigrove offers a "commissive account" of (the illocutionary force of expressed) forgiveness on which the victim commits herself both to the offender's well-being and "to forswearing hostile reactive attitudes and retaliation regarding the forgiven wrongdoing" (2004, 385). On his view, forgiveness is "related" to promising, but the former commitment lacks the binding force of promises and insofar as the victim might wrong the offender in breaching the relevant commitment, there will be an "excusing condition" (ibid, 385-386). Because his sketch of the relevant commitment paints a picture on which its binding force and potential to wrong the offender are somewhat tenuous, I do not include Pettigrove's view among those that I label "strong commitment" views.

<sup>25</sup>See Scarre (2016, 938). Interestingly, though Scarre defends the thesis that forgiveness "once truly granted, cannot be taken back," he allows that forgiveness can be withdrawn in cases where "the forgiving party makes some significant cognitive adjustment to her appraisal of either the offender or the offence, believing that her original forgiveness was granted in error" (2016, 933). He has in mind cases where the offended party judges that she granted forgiveness too lightly, she learns that the offender's repentance was insincere, and new evidence "reveals the original offense in a new light" (2016, 941). If Wanda's remorse wasn't genuine, then on Scarre's view, Victor would be permitted to withdraw his forgiveness. But as I argued above, we needn't presume that her remorse was insincere.

<sup>26</sup>Bennett acknowledges both the strangeness of construing forgiveness as granting the offender authority over her victim and the plausibility of the view that (at least some elements of) forgiveness can be reversed. He writes, "It might seem odd to think that forgiveness should involve giving normative authority to the person who wronged you. However, it is hard to explain why one might be wary of forgiving before one is ready, and how it can happen that one forgives on impulse and then regrets it without appeal to the role of forgiveness in creating new normative bonds. Nevertheless, although these bonds must be robust enough to account for the phenomenon of regret over having forgiven, forgiveness should not be thought of as a once-and-for-all signing away of one's right... Forgiveness does involve binding oneself. However, it is also the case that the wrongdoer can, by a failure in their own redemptive commitments, forfeit their new normative position" (2018, 226-227).

<sup>27</sup>I use this term somewhat reluctantly, as it is not clear to what extent Scarre's and Bennett's respective conceptions of forgiveness incorporate or reflect our ordinary view of full-blown promises. I take it, however, that both accounts stand in tension with a permissive view of un-forgiving because they interpret the type of commitment internal to forgiveness as, in important respects, similar to that of a promise.

There is certainly room for disagreement here about how to construe the strength of the relevant commitments. While I don't want to overstate the demandingness of these normative binds, what I hope to have shown here is that, minimally, their force and structure suggests against a permissive view of justified un-forgiving.

#### 4 | THE NORMATIVE SIGNIFICANCE OF UN-FORGIVING

Above, I articulated a tension between un-forgiving and strong commitment views of forgiveness. Notice that how we choose to navigate this tension might depend on how we regard the normative status of un-forgiving. If, for example, un-forgiving exemplifies a vice or is otherwise morally problematic, then we have added reason to opt for strong commitment views that constrain its justifiability. On the other hand, if un-forgiving is valuable in some respect, then its value might instead support reinterpreting the type of commitment internal to forgiveness in order to accommodate a more permissive view of un-forgiving. Though I offer no definitive resolution to this tension here, in what follows, I make a gestural case for the latter approach.

First, consider the view that un-forgiving sullies or corrupts the value of one's earlier forgiveness. If, as many have suggested, forgiving exemplifies a virtue, then un-forgiving might be viewed as enacting a vice. We ought not to be vindictive or unduly blameful in our relations with others, but we should instead strive to be considerate and understanding toward our fellows. The forgiver (ideally) inhabits the latter stance and in so doing, facilitates what is thought to be a central aim of forgiveness: the maintenance and repair of moral relationships. Accounts that accommodate permissive views of un-forgiving might seem antithetical to this aim.

Unsurprisingly, advocates of strong commitment accounts of forgiveness often tout as an advantage of their views that they cohere well with forgiveness's function of relationship repair. The offender's moral "indebtedness" to her victim serves as a hindrance to the pair engaging amicably as moral equals, and forgiveness removes this barrier and clears a path for (at least partial) relationship restoration. But in order for forgiveness to do its work, we tend to think that it must not represent this change in the normative situation between the victim and her forgiven offender as merely temporary, but as bearing the mark of a kind of closure or finality. Consequently, Bennett insists that in order for forgiveness to allow offenders to be "re-accepted, once and for all," it must involve not only an acknowledgement that the offender no longer stands under certain obligations to the victim, but "an element that binds the parties not to keep on re-opening the matter but rather to bracket it, leave it in the past, and move on" (2018, 229). In a similar vein, Scarre maintains that forgiveness works in the healing of relationships because it is understood "to involve a commitment that will not normally be broken or retracted," where the relevant commitment is, in part and ideally, to let both the offense and the forgiveness itself to "sink into the past" (2016, 943). For these theorists, their accounts fit well with forgiveness's central aim of relationship repair precisely because they imply a restrictive view of un-forgiving.

There is much to say about this. First, while some might seek forgiveness in order to obtain relief from amendatory burdens, or again, to secure from their victims a promise to refrain from future blame, many paradigmatic cases of forgiveness lack this flavor. Consider, for example, that a victim might reasonably be offended if she learned that her offender asked for forgiveness specifically in order to be released from certain obligations or to place her, the *victim*, under a new obligation. Oftentimes, what we seek in asking for forgiveness is not debt-relief or a promise from those we have wronged, but a particular kind of affective reconciliation whereby our victims come to see us, or to hold us, in an improved light.

This is not to say that forgiveness is a matter of mere feelings; rather, it involves a normatively rich form of engagement. We might characterize forgiveness as a mode of valuing or a reason-sensitive, affective holding.<sup>28</sup> Think here of related phenomena, such as moral praise, trust, and love. They each are emotional in nature, but when we praise, trust, or love, we do not just feel a certain way, but we also do a certain thing – i.e., we engage with value. We can add that in praising, trusting, or loving others, we also do more than merely recognize them as praiseworthy, trustworthy, or loveable, but we relate to them in ways that take up the relevant value and lend it a special significance in our interactions. Notice, too, that praise, trust, and love can help to sustain and repair relationships without any guarantee of their continuation. Understanding forgiveness in this way allows us to see how a more permissive view of justified un-forgiving can aid rather than hinder the progression of moral relationships. Let me explain.

In order to illuminate the significance of un-forgiving, it will be helpful to situate it within a broader moral context. Experiences of performing, being the recipient of, or even merely bearing witness to, a wrongful act – or again, an especially kind one – set the stage for the pattern of responses that help constitute our moral relationships. I take it that these responses are modes of valuing that are often historical and dialogic, both familiar points in the praise and blame literature.<sup>29</sup> In other words, they engage with the moral history of particular relationships and shape their future development by consisting in and prompting “replies” to various actions and attitudes. These responses include, among others, blaming and praising attitudes, forgiving, and as I have urged here, *un-forgiving*. To see this, suppose that I wrong my friend, she responds with resentment, and I, with guilt and apology. She then forgives me, and I respond with gratitude. One might think the case closed, but she or I might later perform some act that re-opens this dialogue and creates intersecting ones. And as a result, she may then un-forgive me. Such responses thread the binding of our interpersonal histories, allowing us to navigate (and expand) the narrative dimensions of moral relationships and the shared values that constitute them. But they can only do so if we allow ourselves to engage sufficiently with the past.

This makes it clear why we needn't view un-forgiving as enacting a vice. While I would not deny that un-forgiving can be done in vicious ways and for vicious reasons, it seems to me that forgiveness can go wrong in similar respects.<sup>30</sup> Being too quick to forgive, for example, might show a lack of respect for oneself. A forgiver might also wrong one's offender by forgiving her in ways and under circumstances that she could reasonably reject.<sup>31</sup> In other words, it is not always virtuous to forgive. And to this, I would add that insofar as forgiveness reflects a virtue, being disposed to *un-forgive* – at the right time, in the right way, and for the right reasons – might be partly constitutive of the very same one. Un-forgiving, when done well, tracks the appropriateness of blame and allows relationships to

---

<sup>28</sup>See, for example, Lucy Allais's illuminating discussion of the role of reactive attitudes in forgiving and the rich notion of affective regard that she employs to explain forgiveness (2008, esp. 52-58). See also her defense of affective accounts of forgiveness against rival views (Allais, forthcoming).

<sup>29</sup>This view, or something like it, is pervasive throughout the reactive attitudes literature on moral address. For more on the communicative nature of reactive attitudes, see Walker (2006), Darwall (2006), Shoemaker (2007), Watson (2008; 2012), Strawson (2008), McGeer (2012), McKenna (2012), Helm (2012), and Macnamara (2013) and (2015). For an especially illuminating and persuasive discussion of how forgiveness, and the reactive attitudes more broadly, build and sustain moral community via their roles in “trajectories of reactive exchanges,” see McGeer (2012).

<sup>30</sup>As David Novitz writes, “...to be too willing to undertake the task [of forgiveness], or to undertake it in inappropriate circumstances, is a vice since it is indicative of diminished self-respect.” (1998, 299). See also Garrard and McNaughton for discussion of “objectionably facile forgiveness” (2011, 97).

<sup>31</sup>See, for example Beglin (forthcoming) for a rich and insightful treatment of this phenomenon.

evolve while honestly and reflectively integrating the significance of past moral interactions into their stories.

To my mind, then, we have good reason to afford un-forgiving a substantive role in our moral practices. This needn't mean countenancing capricious un-forgiving, much less endless cycles of extending and withdrawing forgiveness, as acceptable modes of relating to others. But it does favor reinterpreting what is normatively suspect about un-forgiving (when it is) in ways other than casting it as the exercise of a moral vice or, for that matter, the breach of a strong commitment to the wrongdoer.

As a first step toward this end, we should acknowledge that as Brandon Warmke aptly puts the point, having heard that the victim had forgiven an offender, we would expect that the victim "would – and should – no longer blame [her offender]" (2016b, 691). In other words, un-forgiving tends to strike us as a normatively suspect practice. But while strong commitment views of forgiveness offer one helpful, possible explanation in support of this intuition, they do not offer the *only* possible explanation.<sup>32</sup> In closing, I will gesture at an alternative view that captures some of the same intuitions, one that employs resources from the affective-dialogical approach to understanding forgiveness that I outlined above.

Let's begin with cases of morally impermissible un-forgiving. When one purports to forgive an offender for an act and later blames her again for the very same wrongdoing, the offender will often have legitimate cause for complaint. In the first place, the offender might reasonably suspect that the victim culpably misled her in feigning forgiveness rather than truly offering it. And even if she remains convinced that the victim forgave her sincerely, the offender might be concerned that her efforts at amends and reformation were not properly valued. Say, for example, the offender is a close relationship partner whose moral infraction was relatively minor, and she was appropriately remorseful and apologetic, doing more than one could reasonably expect to make up for her wrongdoing. It may be that in such a case, the victim *ought* to no longer blame her.<sup>33</sup> Blame in these circumstances may indicate a moral failure on the part of the victim. But post-forgiveness blame needn't always be (even *prima facie*) morally impermissible.

Suppose that the victim sincerely forgave her offender, that (as I suspect is true in most cases) she was not morally required to do so, and that she later un-forgave her offender, blaming her again for the

---

<sup>32</sup>Though I do not offer a sustained critique of strong commitment views of forgiveness here, I do think that there are independent reasons to reject such views. Here, I will mention just one: they often seem at odds with the phenomenology of ordinary cases of forgiveness. Suppose, for example, that a stranger carelessly injures me, apologizes, and abruptly leaves the vicinity before I have a chance to respond. I might well forgive her without ever expressing it. In such a case, it seems unlikely that I would construe my forgiveness as dissolving a debt, relinquishing a right to blame, or otherwise committing myself to refrain from future blame on pain of doing something morally wrong. One might think that the *expression* of forgiveness is what would constitute the relevant commitment (implicit or otherwise), but it is unclear why the mere utterance, "I forgive you," would add this much. While communicated forgiveness grounds some predictive-cum-normative expectations for future blame, we might do better to avoid needlessly cumbersome interpretations of seemingly simple cases.

The limitations of promissory models of forgiveness seem especially glaring in such cases. In some relationships, it might make sense to supplement one's forgiveness with a commitment akin to a promise, but the types of relationships in which we forgive by far outstrip the sort for which such commitments are apt concomitants. For example, it seems onerous to insist that in forgiving the stranger, I obligate myself to her in the ways that promising suggests. Consider, too, that typically, an agent's promisee has the power to release her from the obligations created by a promise, rendering failure to act in accordance with it permissible where it would not have been so otherwise. But notice that whatever strikes us as normatively fishy about taking back forgiveness would not be alleviated by the offender's assurance, "It's okay – I gave my victim permission to blame me again." The offender's pronouncement here seems not only bizarre and misplaced, but also normatively impotent.

<sup>33</sup>On my view, forgiveness is sometimes morally required, and I suspect that in these cases, it will rarely be permissible to un-forgive. I say rarely (rather than never), because I don't want to rule out the possibility that there might arise a new circumstance or consideration that changes the normative status of the victim's relation to the offender's earlier wrongdoing.

same act. We might understand her blame as normatively suspect without assuming that she is illicitly reinstating a previously cancelled debt, immorally feeling and behaving outside the boundaries of her rights, or breaching a promissory commitment. When we report our forgiveness to the offender, we ground a reasonable – if defeasible – expectation that we will not later resent her for the relevant wrongdoing. It is a commonplace in moral life that forgiveness is not usually reversed. The relevant non-adversarial stance is typically a stable one that persists and often strengthens over time. Having grounded such an expectation, then, a victim who un-forgives at least owes her offender an *explanation* for why her (unrepudiated) blame has returned. Otherwise, the offender would be understandably confused and frustrated. Notice, though, that this explanatory burden is born not of the forgiver's commitment to refrain from future blame, but of the norms of moral communication. In expressing forgiveness, the victim communicates to the offender that she no longer blames him (at least in the same way) for the offense. The offender might reasonably rely on this information in navigating the relationship and without further explanation, would likely find expressed blame for the previously forgiven wrongdoing unintelligible. The victim who expresses forgiveness, then, faces an onus to explain any renewed blame to her offender, qua moral interlocutor, so that the offender can revise his expectations accordingly and make sense of their relationship. We have reason to think that instances of un-forgiving that run afoul of this requirement are normatively problematic.<sup>34</sup>

Post-forgiveness blame may be normatively suspect in still another sense, and this applies both to expressed forgiveness and to unexpressed, or private, forgiveness. It is not mere happenstance that forgiveness is not usually reversed. One way to explain the relative stability of forgiveness is the forger's strong commitment not to blame in the future, but as I have been arguing, we needn't construe forgiveness's commitment(s) in those terms. Another, and I think, more promising route to explaining the relative stability of forgiveness appeals to the reasons for which we forgive. Recall that the victim will often have reasons that justify without rationally requiring her forgiveness. If the victim initially saw sufficient reason to cease blaming the offender, then *ceteris paribus*, her initial forgiveness should exert some rational pressure on her to be consistent, and so, not to reverse her forgiveness absent some new development. But we must not put the point too sharply. The offender's wrongdoing justified her initial blame, and that wrongdoing was not undone. As such, it, along with other new developments, would factor into the victim's reasons for un-forgiving. This may be the best way to describe our earlier scenario in which Victor un-forgives Wanda. Wanda may be within her rights to ask Victor for an explanation of his negative shift in attitude toward her, but if he has un-forgiven her with good reason, then further protest on her part will likely be unwarranted.<sup>35</sup>

The point here is that we don't need strong commitment views of forgiveness to explain the sense in which un-forgiving (often) seems normatively suspect. Understanding the forger as bound by certain norms of moral communication and rationality can do some explanatory work in this regard. In typical cases, the victim who expressed forgiveness should explain – or better, justify – to her offender any later instance of un-forgiving. The forger who refuses to do this, *ceteris paribus*, does something normatively suspect. Relatedly, the rational norms of forgiveness will not countenance capricious un-forgiving, and the forger who engages in this practice will be criticizable. This picture allows for a more permissive, though not entirely normatively unconstrained, view of un-forgiving.

---

<sup>34</sup>I do not mean that a victim who forgives, and then later un-forgives, a stranger is required to track her offender down and explain the change. This norm applies only in cases where future communication is likely or the offender is otherwise likely to have reason to rely on her victim's earlier expressed forgiveness.

<sup>35</sup>This is not to say that Wanda wouldn't still *feel* slighted and continue to protest Victor's blame. She might insist that despite her recent attitude and actions, Victor nonetheless wrongs her in blaming her now for her previously forgiven affair. Of course, on the view that I advocate for here, Wanda would be mistaken.

Doubtless, some will find this way of cashing out the normative significance of un-forgiving unsatisfying and prefer instead those approaches that paint un-forgiving with a more morally critical brush. But for those who are inclined toward a more flexible role for un-forgiving in our moral practices, this approach might constitute an attractive alternative.

## 5 | CONCLUSION

In this paper, I argued that we sometimes un-forgive justifiably, articulated a tension between (what I have called) strong commitment views of forgiveness and a permissive view of justified un-forgiving, and sketched an approach to understanding the normative significance of un-forgiving that favors a more permissive view. Considering how we might un-forgive for good reasons throws into stark relief the need to scrutinize forgiveness's common association with a firm commitment to closure. Similarly, taking un-forgiving seriously challenges us to examine its normative significance through a softer lens, one that can accommodate a view on which, un-forgiving – while sometimes problematic – might nonetheless play an important, positive role in our moral practices.

To be sure, a permissive view of un-forgiving does not cohere well with some conceptions of how forgiveness “wipes the slate clean.” As I noted in the introduction, clean slates are often hard to come by, and I will now add that they may be overrated. A slate's markings may tell a deep and interesting story about its history. As Aurel Kolnai insightfully remarks, “A serious wound when healed leaves a scar behind; it cannot vanish to the point of *restitutio in integrum*. On the other hand, in some sense the ‘scar’ may prove to be tougher, more solid and resistant, than the intact tissue was” (1974, 102). Our relationships with others render us vulnerable to damage, but in at least some cases, it might be better to own, rather than erase, the injuries that we've weathered in order to maintain those relationships – and if necessary, be prepared to face those injuries again.<sup>36</sup>

## REFERENCES

- Allais, L. (2008). Wiping the slate clean: The heart of forgiveness. *Philosophy and Public Affairs*, 36, 33–68. <https://doi.org/10.1111/j.1088-4963.2008.00123.x>
- Allais, L. (forthcoming). Feeling blame and feeling forgiveness. In G. Pettigrove & R. Enright (Eds.), *Routledge handbook of forgiveness*. New York and London: Routledge.
- Armstrong, B. J., Caffey, C., & Wiedlin, J. (2001). *Unforgiven. God bless the Go-Go's*. Record label: Beyond.
- Beglin, D. (forthcoming). Unconditional forgiveness and normative condescension. In D. Shoemaker (Ed.), *Oxford studies in agency and responsibility* (Vol. 7). New York, NY: Oxford University Press.
- Bell, M. (2008). Forgiving someone for who they are (and not just what they've done). *Philosophy and Phenomenological Research*, 77, 625–658. <https://doi.org/10.1111/j.1933-1592.2008.00213.x>
- Bennett, C. (2003). Personal and redemptive forgiveness. *European Journal of Philosophy*, 11, 127–144. <https://doi.org/10.1111/1468-0378.00179>
- Bennett, C. (2018). The alteration thesis: Forgiveness as a normative power. *Philosophy and Public Affairs*, 46, 207–33. <https://doi.org/10.1111/papa.12117>

<sup>36</sup>Earlier versions of this paper were presented at the Princeton University Center for Human Values, the Center for Values and Social Policy at the University of Colorado at Boulder, the 2017 American Philosophical Association Pacific Division Meeting, the Zicklin Center Normative Business Ethics Workshop at the Wharton School at the University of Pennsylvania, UC Riverside, UC San Diego, and UCLA. I thank the audiences at these events, as well as Craig Agule, Lucy Allais, David Beglin, Mitch Berman, Gunnar Björnsson, Jeffrey Blustein, Rosalind Chaplin, Myisha Cherry, John Doris, John Martin Fischer, Elizabeth Harman, Pamela Hieronymi, Agnieszka Jaworska, Karen Jones, Coleen Macnamara, Victoria McGeer, Dana Nelkin, Hanna Pickard, Eric Schwitzgebel, Amy Sepinwall, Julie Tannenbaum, Manuel Vargas, and an anonymous referee from *Philosophy and Phenomenological Research* for helpful comments and discussion.

- Blustein, J. M. (2014). *Forgiveness and remembrance: Remembering wrongdoing in personal and public life*. Oxford: Oxford University Press.
- Bovens, L. (2008). Apologies. *Proceedings of the Aristotelian Society*, 108, 219–239. <https://doi.org/10.1111/j.1467-9264.2008.00244.x>
- Bovens, L. (2009). Must I be forgiven? *Analysis*, 69, 227–233. <https://doi.org/10.1093/analys/69/3>
- Butler, J. (1897). *The Works of Joseph Butler*, Vol. 2, *Fifteen Sermons* [1726], W. E. Gladstone (Ed.). Oxford: Clarendon Press.
- Calhoun, C. (1992). Changing one's heart. *Ethics*, 103, 76–96. <https://doi.org/10.1086/293471>
- Cornell, N. (2017). The possibility of preemptive forgiving. *Philosophical Review*, 126, 242–272. <https://doi.org/10.1215/00318108-3772018>
- Garcia, E. (2011). Bishop Butler on forgiveness and resentment. *Philosophers' Imprint*, 11, 1–19.
- Garrard, E., & McNaughton, D. (2011). Conditional unconditional forgiveness. In C. Fricke (Ed.), *The ethics of forgiveness* (pp. 97–106). New York: Routledge.
- Griswold, C. L. (2007). *Forgiveness: A philosophical exploration*. Cambridge: Cambridge University Press.
- Hellenga, R. (2016). The truth about death. In R. Hellenga (Ed.), *The truth about death and other stories* (pp. 1–140). New York: Bloomsbury, Kindle.
- Hieronymi, P. (2001). Articulating an uncompromising forgiveness. *Philosophy and Phenomenological Research*, 62, 529–555. <https://doi.org/10.1111/j.1933-1592.2001.tb00073.x>
- Kolnai, A. (1974). Forgiveness. *Proceedings of the Aristotelian Society*, 74, 91–106.
- Martin, A. M. (2010). Owning up and lowering down: The power of apology. *Journal of Philosophy*, 107, 534–553. <https://doi.org/10.5840/jphil20101071037>
- Milam, P.-E. (2019). Reasons to forgive. *Analysis*, 79, 242–251. <https://doi.org/10.1093/analys/any017>
- Murphy, J. G., & Hampton, J. (1988). *Forgiveness and mercy*. Cambridge: Cambridge University Press.
- Nelkin, D. (2013). Freedom and forgiveness. In I. Haji & J. Caouette (Eds.), *Free will and moral responsibility* (pp. 165–188). Newcastle, UK: Cambridge Scholars Publishing.
- Novitz, D. (1998). Forgiveness and self-respect. *Philosophy and Phenomenological Research*, 58, 299–315. <https://doi.org/10.2307/2653510>
- Owens, D. (2012). *Shaping the normative landscape*. Oxford: Oxford University Press.
- Pettigrove, G. (2004). The forgiveness we speak: The illocutionary force of forgiving. *Southern Journal of Philosophy*, 42, 371–392. <https://doi.org/10.1111/j.2041-6962.2004.tb01938.x>
- Richards, N. (1988). Forgiveness. *Ethics*, 99, 77–97. <https://doi.org/10.1086/293036>
- Roberts, R. C. (1995). Forgivingness. *American Philosophical Quarterly*, 32, 289–306.
- Russell, L. (2016). Forgiving while punishing. *Australasian Journal of Philosophy*, 94, 704–718. <https://doi.org/10.1080/00048402.2016.1147056>
- Scarre, G. (2016). On taking back forgiveness. *Ethical Theory and Moral Practice*, 19, 931–944. <https://doi.org/10.1007/s10677-015-9651-z>
- Twambley, P. (1976). Mercy and forgiveness. *Analysis*, 36, 84–90. <https://doi.org/10.1093/analys/36.2.84>
- Warmke, B. (2016a). The economic model of forgiveness. *Pacific Philosophical Quarterly*, 97, 570–589. <https://doi.org/10.1111/papq.12055>
- Warmke, B. (2016b). The normative significance of forgiveness. *Australasian Journal of Philosophy*, 94, 687–703. <https://doi.org/10.1080/00048402.2015.1126850>
- Warmke, B., & McKenna, M. (2013). Moral responsibility, forgiveness, and conversation. In I. Haji & J. Caouette (Eds.), *Free will and moral responsibility* (pp. 189–211). Newcastle, UK: Cambridge Scholars Publishing.
- Wonderly, M. (forthcoming). Can we un-forgive? *Philosophers' Imprint*.

**How to cite this article:** Wonderly M. Forgiving, Committing, and Un-forgiving. *Philo Phenomenol Res.* 2021;00:1–15. <https://doi.org/10.1111/phpr.12772>