



The Normative Significance of Forgiveness

Brandon Warmke

University of Notre Dame

ABSTRACT

P.F. Strawson claimed that forgiveness is such an essential part of our moral practices that we could not extricate it from our form of life even if we so desired. But what is it about forgiveness that would make it such a central feature of our moral experience? In this paper, I suggest that the answer has to do with what I will call the *normative significance* of forgiveness. Forgiveness is normatively significant in the sense that, in its paradigmatic instances, forgiving alters the operative norms bearing on the interaction between the victim and the wrongdoer in certain characteristic ways. My project here is, first, to clarify the ways that paradigmatic cases of forgiveness alter the norms of interaction between victim and wrongdoer and to argue that it is in this respect that forgiveness is a normatively significant feature of our moral responsibility practices. Second, I show that most extant theories of forgiveness fail to explain the characteristic ways in which forgiving alters norms. Third, I offer a theory of forgiveness that accounts for this significant normative feature. I conclude by addressing two objections to my proposal.

ARTICLE HISTORY Received 16 April 2015; Revised manuscript received 9 November 2015

KEYWORDS forgiveness; blame; moral responsibility; P.F. Strawson; normative power

1. Introduction

In his landmark paper 'Freedom and Resentment', P.F. Strawson observed with apparent lament that the topic of forgiveness had become 'a rather unfashionable subject in moral philosophy' [1962: 76]. As Strawson saw things, forgiveness is such an essential part of our moral practices that we could not extricate it from our form of life even if we so desired [ibid.: 80]. That forgiveness had received so little philosophical attention revealed a significant oversight.¹ Yet while, for whatever reason, forgiveness had been unfashionable in the 1960s, philosophical interest in the topic has surged in recent years.² In a steady stream of work, philosophers have developed an impressive array of accounts of both the nature and the ethics of forgiveness. Left unexplored, though, is the question of why Strawson thought that forgiveness is so vital to our moral life as we know it. Why is forgiveness such a central feature of our moral experience?

¹ *The Philosopher's Index* reveals only two entries from 1940–1962 with 'forgiveness' in the title, both of them on theological topics. Strawson wasn't alone in noticing how unpopular the topic of forgiveness had become. In 1943, C.S. Lewis went so far as to claim that the virtue of forgiveness was less popular than chastity [104].

² Anthony Bash observes that a search of *The Philosopher's Index* for work on forgiveness 'reveals 269 entries for the period of 1940–2005, of which 132 entries (almost half) were published in the period 2000–2005' [2007: 31]. The *Index* records another 100 entries from 2006–2013 of works with 'forgiveness' in the title.

The answer is that forgiveness is *normatively significant*. The manner in which forgiveness is so will become clear shortly, but the general idea is that forgiving alters the norms of interaction between victim and wrongdoer. By this, I mean that forgiving affects the operative standards governing how the victim and wrongdoer are morally obliged or permitted to regard and treat one another. After forgiving the wrongdoer, for example, it is no longer appropriate for the victim to treat the wrongdoer in certain ways (ways constitutive of blaming behaviours, perhaps). As I will show, this norm-alteration is uniquely accomplished by forgiveness and is so embedded in our moral practice that we rarely notice it—yet it is at once obvious when brought to our attention.

The problem is that most extant theories of forgiveness lack the resources to explain how forgiveness alters these norms of interaction, something that any plausible theory of forgiveness should be able to explain. My project here is to bring into relief the ways in which forgiving alters norms, and to put forward a theory of forgiveness that makes sense of its normative significance.

I proceed as follows. First, I clarify the ways in which paradigmatic cases of forgiveness alter the norms of interaction between victim and wrongdoer. It is in this respect that forgiveness is a normatively significant feature of our moral responsibility practices. This is why it makes sense to hold—like Strawson—that forgiveness is inextricably linked to our form of life. Second, I show that most extant theories of forgiveness fail to explain why forgiving alters norms and that this is a serious problem for such theories. Third, I offer a theory of forgiveness that accounts for this normative significance. I conclude by addressing two objections.

2. Forgiving Alters Norms

Consider the following case:

DECEPTION. Julia and Otto are friends and neighbours. Julia has told Otto that she will pick up Otto's daughter from school. Right before she needs to leave to pick up his daughter, she decides that she doesn't really feel like doing so, and phones Otto to tell him that she can't do so. This causes Otto to be an hour late in picking up his daughter. The next week, Julia feels terrible for deceiving Otto, visits him to confess, apologize, and tells Otto that she is willing to pick up his daughter from school for the next two weeks. Otto accepts Julia's apology, and tells Julia that he understands because we all do selfish things from time to time, and that he forgives her. They reconcile their friendship and continue on with their relationship.

I ask the reader to assume three things about this case: (1) that Julia is responsible and blameworthy for wronging Otto; (2) that it was permissible for Otto to blame Julia in certain ways; and (3) that Otto forgave Julia (according to whatever account of forgiveness the reader prefers) and that his forgiveness was morally permissible.

Now suppose that, the next day, Otto approaches Julia with resentment, demands an apology for deceiving him, tells her what a shoddy thing she did and that she is obliged to give his daughter rides home from school for the next week. It is, I think, obvious that Otto has done something *prima facie* morally impermissible by regarding and treating Julia in these ways when he has forgiven her for deceiving him.³ This feature of

³ I say that Otto has done something *prima facie* morally impermissible because there may be an all-things-considered obligation to continue to blame if, say, not blaming would cause the death of innocent children. In the absence of other considerations, it is typically wrong to keep blaming the wrongdoer after forgiving her. I will suppress this caveat.

Otto and Julia's story generalizes: upon forgiving, some of what were previously permissible modes of regarding and treating a wrongdoer become impermissible. Indeed, it is *because* Otto forgave Julia that certain modes of interaction that *had been* permissible are made impermissible. It would be natural for a third-party observer to think, 'I thought that Otto had forgiven her; why is he treating her like that?'⁴

In paradigmatic cases, forgiveness alters the *victim norms*, the operative norms that govern how the victim interacts with the wrongdoer. What alterations are made to the victim norms? I submit that, when we are the victims of interpersonal wrongdoing, we thereby inherit certain sorts of rights or permissions—namely, the right or permission to *blame* the wrongdoer in certain ways.⁵ Some of these blaming-rights are shared by others in the moral community: rights to censure or denounce, for example, are often held by third parties. Other blaming-rights are usually unique to the victim—such as the right to request an apology or demand restitution, or the right to alter the terms of that specific relationship in certain ways (such as to withdraw from weekly lunches). When a victim forgives his wrongdoer, he relinquishes at least some of these rights. So, one way to explain why Otto should not request further apologies, demand restitution, or censure or denounce Julia is that, by forgiving her, he has given up his right to do so.⁶

It might be thought that, by forgiving, the victim also relinquishes certain rights to *regard* the wrongdoer in certain ways. Paul Twambley [1976: 89] claims that, when you are wronged by another, '[y]ou are within your right to resent his action. In forgiving him, you relinquish that right, you readjust your relationship to one of equality.' Perhaps, then, the forgiver also relinquishes the right to manifest and embrace certain kinds of private blaming attitudes, such as resentment and rancour (or perhaps even sadness, disappointment, and hurt feelings).⁷

There is no easy way to draw the line between the interactions that become impermissible due to forgiveness and those that do not. Where the line is drawn in any given case will likely depend on contextual features: the nature of the relationship, the seriousness of the wrong committed, and the sincerity of the wrongdoer's apology and change of heart, among other things. What is important is that, wherever the line is drawn, there *is* a line: some previously permissible modes of treating the wrongdoer become impermissible upon forgiving.

And yet paradigmatic cases of forgiveness also alter the operative norms governing the wrongdoer's interactions with the victim—the *wrongdoer norms*. Consider Dana Nelkin's [2013] suggestion that forgiveness can involve the release of a special kind of

⁴ One might object that Otto's later treatment of Julia shows that he never forgave her in the first place. It might be true that Otto never forgave Julia, but this would not change the fact that *if* he had forgiven her, then those post-forgiveness interactions would still be inappropriate (regardless of whether or not he engaged in them).

⁵ The matter is complicated by sceptics (e.g. Pereboom [2001, 2014]) who argue that we are not morally responsible and blameworthy (in a certain desert-entailing accountability sense) for our conduct. If this were so, then victims could not have the right to blame the wrongdoer in the first place, and so would not be able to relinquish such rights. However, sceptics could argue that victims relinquish other rights, like the right to protest the wrongdoing or to use the wrongdoing as a reason to weaken or dissolve the relationship (cf. Pereboom [2014: 188–9]). I am sceptical that the responsibility sceptic can retain a distinct and substantive notion of forgiveness, but that is a discussion for another time (see Capes and Warmke [manuscript]).

⁶ I speak about relinquishing rights or permissions to blame, but I am open to other characterizations of the changes in victim norms.

⁷ It is traditional to understand the relevant emotions narrowly: resentment, indignation, or disapprobation (see, e.g., Hughes [1993]). I mention sadness, disappointment, and hurt feelings because of the recent movement to expand the class of emotions that may or must be given up when one forgives (see, e.g., Richards [1988], Murphy [2003: 59], and Blustein [2014: ch.1]). I will not take a stand on this matter here.

personal obligation that the wrongdoer has to her victim. On this view, wrongdoers owe something to their victims. (This is, I suspect, why we sometimes speak of wrongdoers as having a ‘debt to pay’.) What a wrongdoer owes to her victim will often depend on the context, but will generally include sincere apology, repentance and remorse, restitution (which seeks to compensate the wrongdoer for the harm done), and penance (which goes *beyond* mere compensation and makes a costly sacrificial gift to the victim).⁸ In forgiving, the victim can release the wrongdoer from certain personal obligations to do these things.

The obligations are *personal* because the wrongdoer can be released only from those obligations over which the victim has normative authority. For example, a wrongdoer might have obligations to apologize or make amends due to reasons having to do with etiquette, the law, or one’s personal moral commitments. The victim does not have the normative standing to release the wrongdoer from these obligations. But it still makes perfect sense for a victim to tell the wrongdoer, ‘Look, as far as our relationship is concerned, you are forgiven: you don’t need to keep apologizing and I do not expect further restitution or penance.’

Taking all of this together, call the claim that forgiveness typically alters the norms of interaction for both the victim and wrongdoer the *Post-Forgiveness Fact*:

(PFF) Paradigmatic cases of forgiving alter the norms of interaction for both the victim and the wrongdoer in certain characteristic ways.

A few comments are in order. First, I am not claiming that forgiveness *just is* an alteration of the relevant norms. Forgiveness involves much more besides (a point to which we shall return below). But, no matter what else is involved, our best theories should possess the resources to account for the PFF.

Second, the manner in which forgivers alter the relevant norms may fruitfully be thought of as an exercise of a normative power (see, for example, Hart [1964], Raz [1972], and Chang [manuscript]). Promising, for example, is commonly cited an exercise of a normative power. In *The Concept of Law*, H.L.A. Hart explains [1964: 43]:

To promise is to say something which creates an obligation for the promisor: in order that words should have this kind of effect, rules must exist providing that if words are used by appropriate persons on appropriate occasions (i.e. by sane persons understanding their position and free from various sorts of pressure) those who use these words shall be bound to do the things designated by them. So, when we promise, we make use of specified procedures to change our own moral situation by imposing obligations on ourselves and conferring rights on others; in lawyers’ parlance we exercise ‘a power’ conferred by rules to do this.⁹

While differing in significant respects, both forgiving and promising are ways to alter the normative landscape between agents against the backdrop of a shared set of social and moral norms and concepts. In promising, we impose obligations on ourselves (to do what we promise) and we confer rights on others (to claim what we promise). In forgiving, we relinquish certain rights (for example, to blame) and

⁸ For further discussion, see Swinburne [1989], Smith [2008], and Radzik [2009]. While it is an interesting question *why* wrongdoers have these obligations to their victims, I will not take up this question here. But see, especially, Swinburne [1989: 83–5].

⁹ I am grateful to a referee for this journal for suggesting the connection between PFF and normative powers, and for providing this reference to Hart.

we release others from certain personal obligations (for example, to further apology or restitution).

Understanding forgiveness as an exercise of a normative power makes clear that forgiving does not just have the causal *consequence* of altering certain norms. Many activities can cause norms to be altered: my crying in front of a friend might place her under an obligation to comfort me. Exercises of normative powers, however, are intrinsically related to their results—the alteration of norms is the end state of an act of forgiving. The alteration is a *result* of forgiving and not simply a causal consequence of forgiving.¹⁰

Second, I do not claim that the PFF is true of every phenomenon properly called ‘forgiveness’. But it is true of the paradigmatic cases. In support of this thought, suppose that you hear that, while Julia wronged Otto, he has forgiven her. Upon hearing that Otto forgave Julia, you would *assume* that the norms between them had been altered. You would be surprised to see Otto continue to engage in blaming behaviours, expecting that he would—and should—no longer blame Julia, because he had forgiven her. Embedded in our practices of forgiving is the default assumption that forgiving *typically* results in the alteration of certain kinds of norms. But I do think that there are cases of forgiveness that do not alter norms in these characteristic ways. Forgiveness is a diverse and diffuse practice, admitting of a multitude of modes or ways of forgiving.¹¹ But any account of the paradigmatic cases of forgiveness should at least have the resources to explain the PFF.

What makes a case of forgiveness paradigmatic? In my estimation, the clearest and best manifestations of forgiveness will be the ones that capture a diverse set of facts about forgiveness [Warmke and McKenna 2013; Warmke manuscript b]: (1) forgiveness is often a communicative activity that responds to apologies and to requests for forgiveness; (2) we learn about forgiveness at a young age by watching others model it; (3) forgiveness is often governed by certain kinds of ought-norms, and so is typically under our voluntary control [Warmke forthcoming a]; (4) forgiveness has important social functions and has evolved in most moral communities [McCullough 2008]; (5) forgiveness often functions to let wrongdoers know how they can reasonably expect to be treated by their victim; and (6) forgiveness characteristically alters the norms between victims and wrongdoers.

I argue elsewhere [manuscript b] that, generally speaking, the paradigmatic cases are ones that involve both a mesh or harmony between certain psychological changes and a certain kind of communicative act that is motivated and rationalized by those psychological changes.¹² Regardless of the prospects for that view, all I defend here is the claim that our best accounts of forgiveness should have the resources to explain how forgiving alters norms, even if there are some cases of forgiveness that do not affect the norms in these characteristic ways.

¹⁰ Raz [1972: 84] emphasizes that norms are altered as a *result* of an exercise of a normative power and not simply as a causal *consequence* of it. The result/consequence distinction is Anthony Kenny’s [1968: 15]:

The result of an act is the end state of the change by which the act is defined. When the world changes in a certain way... In that case we may say that the second transformation is a consequence of the first and of the act which brought the first about. The relation between an act and its result is an intrinsic relation, and that between an act and its consequence is a causal relation.

¹¹ See Warmke and McKenna [2013: 198]. Cf. Smith [2008: 134]: ‘notions of forgiveness seem to identify a loose constellation of interrelated meanings among various beliefs, judgments, emotions and actions.’

¹² By the end of the paper, we will have a much better grasp on what I take those changes to be.

We can now see why forgiveness is normatively significant. Once we have a basic system of responsibility practices in place, whereby victims of wrongdoing are permitted to respond to their wrongdoers by blaming them, and whereby wrongdoers have obligations to apologize or give restitution, we have a system in which wrongdoing itself alters interpersonal norms. The practice of forgiveness lets us alter these norms to something resembling an *ex ante* state. This is not something that we could do by justifying or excusing agents for their behaviour (for the victim would not have had the right to blame in the first place). Nor would it be something that we could accomplish by forgetting about what happened or by condoning the behaviour. This unique norm-altering function of forgiveness—this normative significance—explains why someone might think that forgiveness is essential to our moral experience. Forgiveness is inextricably linked to so many other aspects of our moral responsibility practices as we find them: blame, resentment, apology, guilt, and restitution.

3. Extant Theories of Forgiveness

I will now argue that many extant theories of forgiveness do not have the resources to explain how forgiving alters norms. My strategy is to focus on certain commonly cited features of forgiveness and then to investigate whether those features are able to explain the PFF. I should be clear that I do not claim that these features are not central or otherwise constitutive aspects of forgiveness. Rather, I will argue that these features (either individually or in conjunction) cannot explain the characteristic ways in which paradigmatic cases of forgiveness alter norms.

3.1. Overcoming Resentment?

It is widely thought that one forgives by overcoming resentment. Stephen Darwall, for example, tells us that '[t]o forgive is, roughly, to forbear or withdraw resentment' [2006: 72].¹³ Indeed, the view that one forgives by overcoming resentment has so dominated discussions of forgiveness that it has been described as the 'standard view'¹⁴ and as 'received orthodoxy' [Bash 2007: 161].

Can resentment-overcoming views explain the PFF? If so, they must be committed to a claim like this:

Resentment-Overcoming (RO). Forgivers alter the relevant norms of interaction between victim and wrongdoer by overcoming resentment towards the wrongdoer.

RO encounters difficulties. How could giving up my resentment towards someone else have an effect not only on how *I* am morally permitted to treat that someone, but also on how *that person* is morally obliged to treat me? I see no way of linking my overcoming of resentment to the inappropriateness of, say, asking for apologies or engaging in other forms of overt blame. Nor can I see how my overcoming resentment would release you from certain personal obligations to, say, apologize to me or offer me restitution.

There are a number of things that the overcoming of resentment *could* explain. Overcoming my resentment could explain why I am less *motivated* to treat or regard

¹³ Examples include Murphy and Hampton [1988], Hughes [1993], and Holmgren [2012].

¹⁴ See Kekes [2009: 490], Radzik [2009: 117], and Zaibert [2009: 38].

you in certain ways. It could explain why it no longer occurs to me to blame. Overcoming resentment might also explain why I no longer think that there are sufficient reasons to blame. But none of these would explain the altered norms for both parties. Overcoming resentment cannot alter both the victim and wrongdoer norms. RO is false.

If RO is false, this does not mean that forgiving does not require (or does not characteristically involve) the overcoming of resentment.¹⁵ However, if RO is false this *does* mean that, if we want to account for how forgiving characteristically alters certain norms, we cannot hold that paradigmatic forgiveness *just is* the overcoming of resentment.

3.2. Making a Private Commitment?

Perhaps a theory of forgiveness needs something with more normative bite, like a private commitment:

Commitment-Making (CM). Forgivers alter the relevant norms of interaction between victim and wrongdoer by making certain kinds of private commitments to treat and regard the wrongdoer in some ways and not in others.

Such a view claims the following: (1) paradigmatic cases of forgiving involve making certain kinds of private commitments; and (2) these commitments alter the norms of interaction for both the victim and the wrongdoer. Let us take these two claims in turn.

It is not uncommon to find the claim that forgiving characteristically involves making certain private commitments. This is one way to understand those accounts of forgiveness that identify ‘forswearing’ as an important aspect of forgiveness. Strawson claimed that forgiving involves forswearing resentment [1962: 63], and Charles Griswold argues that, for Joseph Butler, to forgive is (in part) to forswear revenge [2007: 36].¹⁶ It may be thought that such acts of forswearing involve a commitment to give up or forgo certain ways of interacting with the wrongdoer.

Supposing that forgiveness typically involves private commitments—what kinds of commitment might be at issue? Glen Pettigrove has suggested two commitments: (1) to forswear hostile reactive attitudes and retaliation toward the wrongdoer; and (2) to treat the wrongdoer with an appropriate level of benevolence [2004: 385]. Let me suggest another: (3) not to treat the wrongdoer in ways that would typically be characterized as blaming behaviours. I suggest this because, while most forms of blame are inappropriate after one has forgiven, such blame need not be hostile or retaliatory, and blame may also be compatible with an appropriate level of benevolence.

It is plausible that paradigmatic cases of forgiveness involve these commitments. But if CM is to explain the PFF, it must show how such commitments alter the relevant norms of interaction. How might commitments do so? Consider the following principle: we should, all else being equal, follow through on our commitments.¹⁷ If I make a New Year’s commitment to exercise more, I should, generally speaking, do so. With

¹⁵ In my view, forgiving does not require overcoming resentment. See, e.g., Garcia [2011], Pettigrove [2012], Nelkin [2013], Warmke and McKenna [2013], Pereboom [2014], and Warmke [manuscript a].

¹⁶ For a different interpretation of Butler’s view, see Garcia [2011].

¹⁷ Margaret Gilbert tells us that, when we make ‘personal commitments’, ‘reason requires one to act in accordance with them, all else being equal’ [2013: 38].

respect to forgiveness, then, if a victim commits not to treat the wrongdoer in ways characteristic of blame, then the victim should, all else being equal, not treat the wrongdoer in these ways. Now we have an explanation for why forgiving characteristically alters the norms of interaction. If the victim privately commits not to blame the wrongdoer, then the victim should not blame her.

There are problems with this strategy. Notice that the discussion so far has focused only on how private commitments alter the *victim norms*. Nothing has been said about how these commitments alter the wrongdoer norms. This lacuna is hardly surprising, for it is difficult to see how the victim's private commitments could release the wrongdoer from personal obligations. Suppose that the victim privately commits to not requesting restitution. This would not imply that the wrongdoer no longer has a personal obligation to offer restitution. Furthermore, it would make little practical sense for the victim to alter the wrongdoer norms if the wrongdoer was not apprised of this fact.

There is a separate question as to whether a private commitment is able to alter even the *victim norms*. The idea motivating CM was that, since forgivers characteristically commit privately not to, say, blame their wrongdoers, this makes impermissible certain modes of previously permissible interaction. I am sceptical of this claim, too, but I will forestall that discussion until a later section. For now, it is enough to note that CM is unable to explain the PFF because it cannot explain how forgiving characteristically alters the wrongdoer norms.

I have argued that neither overcoming resentment nor a private commitment explains why forgiveness characteristically alters the norms of interaction between victim and wrongdoer. I have not addressed the possibility that these two things *jointly* could do so. On this *composite account*, one's overcoming resentment *and* one's private commitment together alter the relevant norms. As an attempt to explain the PFF, though, such an account is unmotivated. What extra resources would be provided by this joint account? I see none.

3.3. A Behabitive Act?

Earlier, I argued that overcoming resentment cannot explain the PFF. Perhaps, however, it is not simply the fact that one overcomes resentment that does the trick, but rather the fact that one *reveals* that one has done so (presumably to the wrongdoer). How does one reveal to one's wrongdoer that one has overcome resentment?

One way is to say to the wrongdoer 'I forgive you' (or to use a cognate expression). In *How to Do Things With Words* [1962], J.L. Austin claimed that there are different ways to understand what we accomplish when we speak. For instance, one way to think of an utterance is as a *locutionary act*, which is simply the act of uttering a sentence with a certain sense and reference. But we do not typically utter sentences simply for the sake of uttering them. In addition to the act *of* uttering a sentence, we may also perform an act *in* uttering a sentence—we ask questions, make demands, warn of threats, persuade detractors, express our preferences, *inter alia*. The things that we accomplish in uttering a sentence, he called *illocutionary acts*. Consider the sentence 'I would like a ham sandwich.' The locutionary act is the utterance of the sentence itself. The illocutionary act might be one of simply asserting a desire, or it might be one of ordering a sandwich at the deli counter.

One way to think about forgiveness, then, is in terms of what we do when we say ‘I forgive you.’ Joram Haber [1991] suggested that ‘I forgive you’ can be understood as a member of a class of illocutionary acts that Austin called *behabitives*, ‘which include the notion of reaction to other people’s behaviour and fortunes and of attitudes and expressions of attitudes to someone else’s past conduct or imminent conduct’ [1962: 159].¹⁸ Examples of *behabitives* include ‘I applaud you’ and ‘I commend you’, which serve to express or exhibit the speaker’s attitude about the conduct of the intended audience. Haber argued [1991: 40] that ‘I forgive you’ can reveal that, while the victim resented that wrongdoer’s action, she has overcome her resentment towards the wrongdoer (or is at least willing to try to overcome it). I think that communicative entities besides utterances could possess this sort of *behabitive* force (for instance, certain kinds of gestures mutually understood to be an expression of a forgiving attitude). So, we need not insist that a speech act occur—just a communicative entity that could function similarly.

Consider, then, another proposal for explaining the PFF:

Behabitive Act (BA). Forgivers alter the relevant norms of interaction between victim and wrongdoer by revealing to the wrongdoer that the forgiver has overcome her resentment towards the wrongdoer.

BA solves one problem encountered above, which was that, because private modes of forgiveness are not communicative entities, it is difficult to see how they could alter the norms in such a way that the wrongdoer could learn of it. But, by *communicating with* the wrongdoer, the victim could alter the wrongdoer norms *and* communicate this change. The problem is that revealing that one has overcome resentment still does not explain how forgiving alters the full range of norms. First, the revelation that one has overcome resentment fails to explain the alterations to the victim norms. If overcoming resentment doesn’t relinquish certain rights to blame, why would revealing this emotional change do so? Further, I cannot see how such a revelation would alter the wrongdoer norms. We are missing a connection between my revealing to you that I no longer feel resentment towards you and your no longer having a personal obligation to, say, apologize or offer restitution to me.

I agree with Haber that sincere utterances of ‘I forgive you’ (as well as paradigmatic cases of forgiveness, more generally) typically possess *behabitive* force. And, as I will argue, understanding paradigmatic manifestations of forgiveness as illocutionary acts does provide the key to explaining the PFF. But we will have to look elsewhere.

I have argued that, even if one reveals to the wrongdoer that one overcomes resentment, this would not explain the PFF. But suppose that we added a private commitment. The result would be a composite account that involves overcoming resentment, revealing this change, and making a private commitment. Would this change matters? Such an account *might* alter some of the victim norms in virtue of the private commitment (although, below, I will raise some doubts about this). But I cannot see how the three features of this composite account would jointly alter the full range of wrongdoer norms. If these three features are singly unable to explain the characteristic ways that paradigmatic cases of forgiveness alter norms, why think that they would do so jointly?

¹⁸ Haber does not claim that uttering ‘I forgive you’ is the only way to forgive; he allows that one can also forgive in other ways [1991: 40].

3.4. A Commissive Act?

Perhaps what we need is a public commitment to the wrongdoer. Building on Haber's work, Glen Pettigrove [2004, 2012] has argued that communicative acts like 'I forgive you' can function as what Austin called a *commissive*. Commissives have the illocutionary force of committing the speaker to a course of conduct. One might say 'I promise to buy you a tangerine tomorrow', or 'I will never lie to you again.' Pettigrove claims that, in addition to functioning as a behabitive, 'I forgive you' can also function as a commissive in exactly this way. Here, too, we need not insist that the commissive force be exerted in a speech act; other communicative entities like gestures or knowing looks will do. So, consider this:

Commissive Act (CA). Forgivers alter the relevant norms of interaction between victim and wrongdoer by making certain kinds of public commitments to the wrongdoer to treat and regard her in some ways and not in others.

If I commit to doing something, then, all else being equal, I should do it. If, in uttering 'I forgive you', I commit to not blaming you, then, all else being equal, I should not do so. Whereas simply revealing that one has given up resentment does not alter the relevant victim norms, a public commitment has the ability to alter what is permissible for the victim. And because CA picks out a communicative act, it is able to explain how the wrongdoer could be informed of the altered norms.

CA comes closer to explaining how forgiving alters norms; but it, too, fails to explain the PFF. First consider whether CA can explain changes to wrongdoer norms. It is hard to see how the relevant commitments could function in this way: your publicly committing not to ask for an apology would not mean that I no longer had a personal obligation to apologize. We can imagine the victim saying, truthfully and felicitously, to the wrongdoer: 'Look, you should still apologize, but I commit not to ask for it.' And so, while CA adds a public dimension that CM lacks, CA is still unable to explain how the wrongdoer norms get altered in the first place. At best, then, CA is capable of explaining only one side of the PFF.¹⁹

But CA falls short here, too. Recall that a commitment is supposed to alter the victim norms because committing to do *x* (or not to do *x*) means that the victim should, all else being equal, do *x* (or not do *x*). But what if the forgiver rescinds her commitment? Does this mean that the forgiver is no longer bound by the new post-forgiveness norms? Commitments are sometimes, after all, a kind of thing that we can take back.

Consider what Margaret Gilbert [2013: 38] calls a 'personal commitment', which is 'a commitment of a person *X*, such that *X* is in a position unilaterally to bring the commitment into being and unilaterally to rescind it.' I might make a personal commitment, via a speech act, to attend your holiday party. Having done so, I should attend the party, all else being equal. But if I rescind this commitment, I am thereby no longer obliged to attend, all else equal. I may be criticizable for breaking my commitment; but I am not usually criticizable for not attending your party, having rescinded my commitment to attend. Suppose that, by forgiving you, I publicly commit to treat and regard you in certain ways and not in others. But then I rescind my commitment. Am I still bound to abide by the norms that have been altered?

¹⁹ A composite account that includes all of the features discussed thus far would still have the problem that it could not explain the characteristic alterations to the wrongdoer norms.

Let us consider two kinds of answer to that question. Here is the first:

(1) After rescinding S's commitment, S is no longer bound by the altered victim norms.

According to (1), what grounds the continued alteration of post-forgiveness norms is the existence of the commitment. Once one retracts the commitment, one is no longer obliged to treat or regard the wrongdoer in ways required by the PFF. (Of course, one might be criticizable for rescinding the commitment itself, but this is a different matter.) The problem with this view is that, in at least some cases, a forgiver is criticizable, not just for rescinding a commitment, but for treating the wrongdoer in ways unbecoming of forgiveness.

In other words, sometimes if I forgive you, then even if I take back the relevant commitment I am still doing something inappropriate by blaming you in various ways. If I rescind my commitment and start blaming you again because I think it would be fun to do so, I can be criticizable for treating you in ways inconsistent with my forgiving you. This is reason to be suspicious of the claim that what undergirds and sustains the alteration of the victim norms is the continued existence of a certain sort of commitment: in at least *some cases*, the norms continue to be altered even when the relevant commitment is rescinded.

A second strategy says that, even if I rescind the commitments at issue when I forgive you, I am still bound by the norm-alterations that were inaugurated by the making of those commitments:

(2) After rescinding S's commitment, S is still bound by the altered victim norms.

This view has the consequence that, while it is a commitment that *inaugurates* the new norms, it is not the commitment's continued existence that *sustains* the norms. If I forgive you and thereby commit not to blame you in certain ways, then even if I rescind my commitment I still should not blame you (all else being equal). But if this is so, it cannot be the current existence of the commitment that explains why I shouldn't blame you after forgiving. Yet we want to know not only what inaugurates the new norms, but why certain modes of treating or regarding you are still impermissible an hour, or a day, or a month after forgiving you. Option (2) cannot explain this because it does not claim that the commitment does the relevant work. Some further explanation would be needed.

This survey of putative explanations of the PFF is not intended to be exhaustive, but it does show how difficult it is for extant theories of forgiveness to account for a common and important aspect of forgiving—the fact that paradigmatic cases of forgiving alter norms for both the victim and the wrongdoer in certain characteristic ways.

4. Forgiveness and Declarative Force

I now suggest a theory of forgiveness that explains the PFF. The basic idea is that paradigmatic cases of forgiveness are communicative acts that possess, not only behabitive and commissive force (*per* Pettigrove), but also declarative force. Declarative acts have the effect of (to put it crudely) changing reality in various ways.²⁰ An appropriate authority (in the right conditions) might say, 'I christen this ship' or 'I hereby find you

²⁰ For more on declaratives, see Searle [1979: 19].

guilty.’ By making such utterances, one is actually able to *make it so* that a ship is christened or that one is found guilty. Understood as a declarative, an overt act (such as an utterance of ‘I forgive you’) makes it the case that one has been forgiven. Such utterances, when made in the right kinds of contexts, bring a state of affairs into existence simply in virtue of it being declared to be so. ‘I forgive you’ would not (just) *commit* the putative forgiver to various things; it would *make it so* that the wrongdoer has been forgiven.

This distinction is key, and seeing how it is employed in different contexts can illuminate the matter. Consider the cancellation of financial debts and judicial pardons. One might commit to treat someone who is a debtor as someone who had no debts, and one might commit to treat a prisoner as someone who was not a prisoner. But in neither case would one be *making it the case that* someone’s debts had been forgiven or that someone had been pardoned. But if someone with the appropriate standing were to *declare* that one’s debts had been forgiven or that one had been pardoned, then (assuming the right kinds of background conditions) this *would* make it so. In the case of forgiving, one might commit to treat the wrongdoer as someone who has been forgiven and yet fail to make it the case that the wrongdoer has been forgiven. In exerting declarative force, however, someone with the appropriate standing can make it the case that a wrongdoer is forgiven. Paradigmatic manifestations of forgiveness, I contend, possess this kind of declarative force.

One reason for thinking that paradigmatic cases of forgiveness possess declarative force is that such force explains how, characteristically, forgiving alters norms. Consider the wrongdoer norms. Upon being forgiven, one is often released from certain kinds of obligation. What one does when one forgives is, in part, to effect this release from such obligations. This is similar to the way in which you release a debtor from the obligation to pay you back, by telling her that she no longer needs to pay you back. Unlike the other proposals, declarative force can explain changes to wrongdoer norms.

Declarative force also explains how the victim norms are altered. Consider again the analogies with economic debt cancellation and pardon. When I cancel your debt by telling you that you no longer need to pay me back, I am not only releasing you from your obligation; I am also relinquishing certain rights, such as the right to demand payment, the right to request interest, and so on. And the governor, when pardoning a criminal, is relinquishing certain rights to keep the criminal in prison. Although I do not want to make too much of these analogies, forgiveness can function in a similar way. Upon declaring the wrongdoer forgiven, one relinquishes certain rights—the rights to demand apology, request restitution, and so on. I therefore suggest the following:

Declarative Act (DA). Forgivers alter the relevant norms of interaction between victim and wrongdoer by declaring the wrongdoer forgiven.

A few clarifications about DA are in order. First, I am not claiming that every instance of forgiveness involves a *speech* act. I submit that acts of forgiveness can have declarative force without being speech acts. In the appropriate context, facial expressions and gestures may serve as a way of forgiving, especially between individuals who share a life together. A simple knowing glance can communicate as much as any spoken word.²¹

²¹ See Swinburne [1989: 85]: ‘Forgiving is a performative act—achieved perhaps by saying solemnly “I forgive you,” or perhaps by saying “That’s all right,” or maybe by just a smile.’

Second, paradigmatic cases of forgiveness do not possess *only* declarative force. Our best accounts of paradigmatic cases of forgiveness will be composite, involving a variety of changes to one's emotions and private commitments. But they will also have an overt communicative dimension with behabitive, commissive, and declarative force. My view is that the paradigmatic cases of forgiveness are those in which there is a *mesh* or *harmony* between the private and overt aspects of forgiveness.²²

Third, I do not claim that all instances of what is properly called forgiveness involve overt behaviours. There are genuine cases of private forgiveness, and such cases will involve some constellation of changes in emotions, revisions in one's judgments, and the formations of new decisions or policies. However, private cases of forgiveness will not involve the full range of normative alterations typically associated with its paradigmatic manifestations. Further, I am open to the possibility that some overt expressions of forgiveness do not possess declarative force. I do not deny that there may be cases of something properly called 'forgiveness' that possess only behabitive force (*per* Haber) or only behabitive and commissive force (*per* Pettigrove).²³

5. Objections to Declarative Force

Some philosophers reject the plausibility of an account of forgiveness that invokes declarative force.²⁴ Here, I discuss two objections raised by Glen Pettigrove.

5.1. The One-Sidedness Objection

For Pettigrove, declarative accounts of forgiveness can be developed in one of three ways [2012: 9–12]. On the 'financial model,' forgiving is like cancelling a financial debt: it cancels or releases someone from moral debts. On the 'juridical model,' forgiving is like pardoning: it releases the wrongdoer from formal or informal sanctions. And on the 'cultic model,' forgiving is like absolution: it wipes the slate clean and removes one's guilt. Pettigrove argues that such views place undue emphasis on the wrongdoer [ibid.: 12]:

[The declarative account] is too focused on the wrongdoer's guilt. The speaker nearly drops out of view. The declarative force of the utterance 'I forgive you' on the financial, juridical, and cultic models would be better served by 'You are forgiven' or 'Your debt is forgiven' or perhaps 'Your sins are forgiven'. The models in which 'I forgive you' has declarative force may come closest to fitting our image of forgiveness when we are in the position of saying, 'Please forgive me'. In those moments we are closest to wanting absolution and most eager to find guarantees that our past failings will not continue to haunt us. These models seem less suitable when we are the one who must offer forgiveness.²⁵

This one-sidedness objection says that, when understood as possessing declarative force, forgiveness is too concerned with the wrongdoer: I cancel *your* debt, I pardon *you*, I wipe away *your* sins. But forgiveness is not one-sided: forgiveness also involves having a change of heart toward the wrongdoer or a commitment to treat and regard the wrongdoer in certain ways from now onwards.

²² I explore this 'mesh' in Warmke [manuscript b].

²³ Compare Neblett [1974: 269]: 'The expression "forgiveness" does not designate a characteristic and peculiar activity; the language of forgiveness is quite often put to a performatory use, and it is not always put to the same performatory use.'

²⁴ Warmke [forthcoming-b] addresses a number of such objections, such as those from Griswold [2007: 61] and Hughes [1993: 337].

²⁵ See also Griswold [2007: 61].

Pettigrove is correct that there are ways of construing declarative force that make forgiveness one-sided. But this is not a reason to reject the claim that forgiveness can possess declarative force. First, even if the most plausible ways of understanding declarative force are one-sided, this does not imply that *forgiveness* itself is similarly one-sided. This is because forgiveness could possess behabitive, commissive, and declarative force. On this view, the forgiver *also* reveals that she had made certain emotional changes and commitments to treat and regard the wrongdoer in certain ways and not others. There is nothing one-sided about this *complete* picture of forgiving. Indeed, Pettigrove himself endorses a similar kind of composite view, holding that forgiveness, when understood as a communicative act, can possess both behabitive and commissive force [2004: 386; 2012: 18–19]. Simply adding declarative force would not make forgiveness one-sided, even if our account of declarative force was one-sided.

Yet we need not construe the declarative force of forgiveness as one-sided in the first place. If forgiveness can alter the norms of interaction for both the victim and the wrongdoer in virtue of its declarative force, then this has consequences for *both* the victim and the wrongdoer. The victim relinquishes certain rights to blame, and the wrongdoer is released from certain kinds of personal obligations. It is difficult to see the manner in which the speaker ‘drops out of view’ on such an account. I agree with Pettigrove that our best accounts of forgiveness should not be one-sided, but I see no good reason to fault either declarative force or accounts of forgiveness that include declarative force.

5.2. The Relationship Objection

In earlier work, Pettigrove raised a different objection: declarative models fail to properly construe forgiveness from the wrongdoer’s perspective [2004: 382]:

But even when we are in the role of penitent wrongdoer the fit between the declarative models and the forgiveness we seek is not exact. Our worry is not that someone or other will come along and prosecute us in moral court. Our anxiety is that *this* person, the one we have wronged, will hold our misdeed against us. Optimally, we are not seeking forgiveness so that each can go his or her own way. We are seeking forgiveness so that we can go on together. The concern is not that I reach a place where I am no longer considered a bad person but that I reach a place where *you whom I have wronged* no longer deem me as such. The models of forgiveness that naturally issue locutions with declarative force do not capture these relational qualities as well as we might hope.

What we want when we are forgiven is not merely to have our names blotted out of the ‘naughty’ book. Rather, we want to know that our victim won’t hold our wrong against us. We want to have the relationship reconciled. We want our victims to regard us as worth caring about. Yet because declarative models of forgiveness focus on changes in formal status, they do not account for the relational aspects of forgiveness. Call this *the relationship objection*.

Pettigrove is correct that there are ways of understanding declarative force that alienate forgiveness from our relationships. But we *need* not think of declarative force in such a way. As I develop the view, forgiveness can effect an alteration of norms of *interaction* between the victim and wrongdoer. On this view, the declarative force of forgiveness does not *merely* alter the formal status of the wrongdoer—it alters the ways in which wrongdoer and victim can and should treat and regard one another. Declarative models of forgiveness therefore need not fall prey to the relationship objection.

Even if that is wrong, this would not mean that any theory of forgiveness that possessed a declarative aspect would thereby fall prey to the relationship objection. If our

accounts of forgiveness can be composite, a plausible account of forgiveness can involve (among other things) *both* commissive and declarative force. By forgiving you, I also commit to regard and treat you in certain ways; and, given that you know that I have so committed, you can have reasonable assurance ‘that we can go on together’. Indeed, on Pettigrove’s own account, this relational aspect of forgiveness is addressed by the commissive force of saying something like ‘I forgive you.’ The resources available to Pettigrove’s own composite account are available to a composite account that includes declarative force. Therefore, even if the declarative force of ‘I forgive you’ is simply a formal change in status (although I have argued that it need not be), there is no good reason to think that a composite account (that includes behabitive, commissive, and declarative force) cannot properly account for the way that forgiveness promotes the reconciliation of a relationship.

6. Conclusion

I began by reflecting on the Strawsonian claim that forgiveness is an essential aspect of our moral responsibility practices as we know them. I have suggested that one way of thinking about this claim is to understand forgiveness as being *normatively significant*, and I have argued for three central claims in order to illuminate the normative significance of forgiveness.

First, in paradigmatic cases of forgiveness, forgiving alters the norms of interaction for both the victim and the wrongdoer in certain characteristic ways: the victim relinquishes certain rights or permissions and the wrongdoer is released from certain personal obligations. Any good theory of forgiveness must have the resources to explain these changes.

Second, I argued that most extant theories of forgiveness do not have these resources. While some views fare better than others, the most popular account of forgiveness—that forgiving is the overcoming of resentment—is unable to explain how forgiveness could alter *either* the victim’s or the wrongdoer’s norms.

Third, I argued that one plausible way of explaining how forgiveness can alter the norms of interaction is by understanding its paradigmatic manifestations as communicative acts that possess (among other things) declarative force: sincerely uttering something like ‘I forgive you’ can have the effect of relinquishing one’s right to blame as well as releasing the wrongdoer from certain personal obligations. On the sort of view I offer, then, paradigmatic instances of forgiveness involve a mesh or harmony between one’s private attitudes towards the wrongdoer and one’s overt acts—acts that (1) reveal that one has undergone certain attitudinal changes, (2) commit the victim henceforth to treating and regarding the wrongdoer in certain ways; and (3) declare the wrongdoer as having a new normative status in relation to the victim, one that alters the norms of interaction between the victim and wrongdoer in certain characteristic ways.²⁶

²⁶ I thank Michael Gill, Michael McKenna, and Mark Timmons for excellent comments on an early version of this paper, and Christian Miller, who carefully read and commented on a later version of it. I also thank two referees and two editors for this journal for the time and attention they gave to my ideas and for helping significantly to improve the paper. Support for work on this paper was funded in part by a grant from the Templeton World Charity Foundation. The opinions expressed in this paper are my own and do not necessarily reflect the views of the foundation.

References

- Austin, J.L. 1962. *How to Do Things With Words*, Cambridge, MA: Harvard University Press.
- Bash, Anthony 2007. *Forgiveness and Christian Ethics*, Cambridge: Cambridge University Press.
- Blustein, Jeffrey M. 2014. *Forgiveness and Remembrance: Remembering Wrongdoing in Personal and Public Life*, New York: Oxford University Press.
- Capes, Justin and Brandon Warmke, [manuscript](#). Forgiveness for Responsibility Sceptics.
- Chang, Ruth, [manuscript](#). Do We Have Normative Powers?
- Darwall, Stephen 2006. *The Second Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- Garcia, Ernesto V. 2011. Bishop Butler on Forgiveness and Resentment, *Philosophers' Imprint* 11/10:1–19.
- Gilbert, Margaret 2013. *Joint Commitment: How We Make the Social World*, New York: Oxford University Press.
- Griswold, Charles L. 2007. *Forgiveness: A Philosophical Exploration*, Cambridge: Cambridge University Press.
- Haber, Joram Graf 1991. *Forgiveness*. Lanham, MD: Rowman and Littlefield.
- Hart, H.L.A. 1964. *The Concept of Law*. Oxford: Clarendon Press.
- Holmgren, Margaret 2012. *Forgiveness and Retribution: Responding to Wrongdoing*, Cambridge: Cambridge University Press.
- Hughes, Paul M. 1993. What Is Involved in Forgiving? *The Journal of Value Inquiry* 27/3–4: 331–40.
- Kekes, John 2009. Blame Versus Forgiveness, *The Monist* 92/4: 488–506.
- Kenny, Anthony 1968. Intention and Purpose in Law, in *Essays in Legal Philosophy*, ed. Robert S. Summers, Berkeley: University of California Press: 146–63.
- Lewis, C.S. 1943 (1996). *Mere Christianity*, New York: Touchstone.
- McCullough, Michael E. 2008. *Beyond Revenge: The Evolution of the Forgiveness Instinct*, San Francisco: Jossey-Bass.
- Murphy, Jeffrie G. 2003. *Getting Even: Forgiveness and Its Limits*, New York: Oxford University Press.
- Murphy, Jeffrie G. and Jean Hampton 1988. *Forgiveness and Mercy*, Cambridge: Cambridge University Press.
- Neblett, William R. 1974. Forgiveness and Ideals, *Mind* 83/330: 269–75.
- Nelkin, Dana 2013. Freedom and Forgiveness, in *Free Will and Moral Responsibility*, ed. Ishtiyaque Haji and Justin Caouette, Newcastle upon Tyne: Cambridge Scholars Press: 165–88.
- Pereboom, Derk 2001. *Living Without Free Will*, Cambridge: Cambridge University Press.
- Pereboom, Derk 2014. *Free Will, Agency, and Meaning in Life*, Oxford: Oxford University Press.
- Pettigrove, Glen 2004. The Forgiveness We Speak: The Illocutionary Force of Forgiving, *The Southern Journal of Philosophy* 42/3: 371–92.
- Pettigrove, Glen 2012. *Forgiveness and Love*, Oxford: Oxford University Press.
- Radzik, Linda 2009. *Making Amends: Atonement in Morality, Law, and Politics*, New York: Oxford University Press.
- Raz, Joseph 1972. Voluntary Obligations and Normative Powers, *Proceedings of the Aristotelian Society* 46/Supplement: 59–102.
- Richards, Norvin 1988. Forgiveness, *Ethics* 99/1: 77–97.
- Searle, John R. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge: Cambridge University Press.
- Smith, Nick 2008. *I Was Wrong: The Meanings of Apologies*, Cambridge: Cambridge University Press.
- Swinburne, Richard 1989. *Responsibility and Atonement*, Oxford: Clarendon Press.
- Strawson, P.F. 1962 (2003). Freedom and Resentment, in *Free Will, 2nd edn*, ed. Gary Watson, Oxford: Oxford University Press: 72–93.
- Twambley, P. 1976. Mercy and Forgiveness, *Analysis* 36/2: 84–90.
- Warmke, Brandon, [forthcoming a](#). Articulate Forgiveness and Normative Constraints, *Canadian Journal of Philosophy*.
- Warmke, Brandon, [forthcoming b](#). The Economic Model of Forgiveness, *Pacific Philosophical Quarterly*.

Warmke, Brandon, [manuscript a](#). Forgiving and Forswearing Resentment.

Warmke, Brandon, [manuscript b](#). Modes of Forgiveness.

Warmke, Brandon and Michael McKenna 2013. Moral Responsibility, Forgiveness, and Conversation, in *Free Will and Moral Responsibility*, ed. Ishtiyaque Haji and Justin Caouette, Newcastle upon Tyne: Cambridge Scholars Press: 189–212.

Zaibert, Leo 2009. The Paradox of Forgiveness, *Journal of Moral Philosophy* 6/3: 365–93.