

## V—FORGIVENESS AND WEAK AGENCY

LAURENT JAFFRO

Forgiveness involves a process, not an isolated act or decision. The initial step lies within the voluntary control of the forgiver. The immediate outcome of the commitment to forgive is the formation of a new context that modifies some of the circumstances for the forgiver as well as for the wrongdoer. Further consequences, notably changes in the forgiver's desires and feelings, cannot be brought about directly. A sound account of forgiveness should focus on its intertemporal structure and highlight the relation between the initial commitment and the subsequent process.

This account of forgiveness is about people who have difficulty controlling their vindictive desires, for reasons either contextual or constitutional. For them, forgiveness is both the problem and the solution: it is problematic because they cannot change their desires at will in order to forgive; however, they can influence their desires over time by deciding to forgive.

Sketching out what constitutes forgiveness is quite tricky, for at least two reasons. First, *forgiveness* is not the name of a single distinct operation, but rather of a collection of practices which vary within and across particular cultures. The notion encompasses institutional as well as personal forgiveness, and is flexible enough to extend into the areas of excuses, clemency, pardon or amnesty, and so on. Second, accounts of forgiveness are shaped by normative considerations. Suppose one thinks that a major reason for *A* to forgive *B* his or her fault is simply that yielding to resentment makes *A* unhappy. This would not lead to including *B*'s recognition of the wrong as a necessary condition. Some, who consider that this would be a bad reason to forgive, would insist that the practice of forgiveness is conditional upon *B*'s recognition, which is required to preserve *A*'s self-respect. Thus, normative considerations affect the structure of forgiveness. However, my question is not about what kind or degree of forgiveness a theory of forgiveness should value

from a moral viewpoint; it is about how to account for how forgiveness functions. I leave aside considerations about moral or prudential reasons, albeit that these are relevant to the understanding of the conditions of forgiveness, and concentrate on its inner workings.

Is forgiving a demonstration of the forgiver's freedom and strength in the face of an agent weak enough to be at fault? There is a common view of the forgiver as a sovereign agent endowed with the power of bestowing pardon, in full control of the operation and its effects. I suggest we turn the tables and look at the forgiver as a weak agent—someone who lacks direct and synchronic control over his or her desires and volitions—and forgiveness as a technique suitable to weak agency.

I will claim that although there is a voluntary dimension to forgiveness that culminates in the decision to forgive, that decision is only a preliminary step, not the whole of forgiving. Forgiveness becomes intelligible when it is envisaged as involving a process, not an isolated act or decision. The initial step lies within the voluntary control of the forgiver. The immediate and intentional outcome of the decision to forgive is the formation of a new context that modifies some of the circumstances for the forgiver as well as for the wrongdoer. Further consequences, notably changes in the forgiver's desires and feelings, which would have been very difficult, if not impossible, to achieve without that modification of context, may be viewed along the lines of what Jon Elster calls states that are essentially by-products (1983, pp. 43–56), that is, upshots that cannot be brought about knowingly and intentionally or at will. The decision to forgive often results in a state of having forgiven, but not in the same way that the decision to stand up and walk away results in my leaving my office. A complex and fallible process, of a causal and mediate nature, occurs between the decision to forgive and successful forgiveness. Thus, a sound account of forgiveness should be dynamic, that is, should focus on its intertemporal structure.

What do we decide when we decide to forgive? The two-stage account claims (1) that we decide to try to forgive, in so far as forgiving is taken in the sense of a process of revision of desires and feelings, and (2) that trying consists in deciding not to act in a vindictive manner, which is also what we mean by 'forgiving'. Thus, we decide to try to forgive in the former sense by deciding to forgive in the latter sense.

In the first part of the paper, I sketch out the basic features that matter to my two-stage account of forgiveness. The second part tries to shed light on the relation between the two stages by describing it in terms of causes and conditions.

## I

*A Dynamic View of Forgiveness.* Forgiving is a three-term operation. It involves the offended as forgiver, the offender, and the offence—a fault serious enough to merit revenge, punishment or forgiveness. In cases of personal forgiveness, nobody but the offended (or perhaps a very close proxy who is also significantly affected by the offence) is qualified to forgive. This is not true in cases of institutional forgiveness, which is often given on behalf of others.

Why three terms? Why do we need to mention both the offence and the offender? Because it is not the offence (the object of forgiveness), but the offender (the addressee) who may be redeemed by forgiving.<sup>1</sup> The transformation to which forgiveness opens the way changes not the fault but the culprit. The offence does not lose its status as a serious wrong. Otherwise, a forgiven offence would not still be liable to other concurrent responses, in particular, punishment. It does make sense to punish someone for a wrong that has been forgiven as long as it is a matter of personal forgiveness (that is, bestowed by the wronged party). The situation is quite different in most cases of political forgiveness, where claimants expect from the institution something distinct from personal revenge: instead of trying to get even with offenders themselves, they ask a third party to punish, so that institutional forgiveness would consist in renouncing, diminishing or suspending punishment. In a legal system in which punishment is sharply distinguished from revenge, institutional forgiveness involves managing punishment, adjusting penalties. Divine forgiveness by a God able to punish and unable to take revenge is on the same footing. On the contrary, in cases of personal forgiveness, the forgiver waives a vindictive course of action; since forgiveness is

<sup>1</sup> Although I have renounced using *forgivee*, coined by some on the pattern of *promisee*, I try to avoid *forgiven* to refer to the person. The logic of forgiveness requires a double object: though it is correct to say that one forgives somebody for doing something, it is nonetheless correct to say that one forgives somebody (in the dative) something (in the accusative). The latter construction parallels *pardonner quelque chose à quelqu'un*, or *jemandem etwas verzeihen*.

supposed to have an effect on personal revenge, nothing prevents it from being combined with punishment. Otherwise it would be identical with mercy (Garrard and McNaughton 2003, p. 48). There is nothing inconsistent in stating ‘I forgive him but I want him to be punished’, whereas it would be inconsistent to say ‘I forgive him but I will get even with him’.

Private forgiveness and legal punishment can go hand in hand. Their compatibility depends on distinctions between the roles of victims and the courts, and the fact that ‘wiping the slate clean’ does not mean erasing the wrong (Allais 2008), but rather closing a painful chapter and perhaps moving forward. Forgiveness is a technique of indirect action on the wrongdoer and the wronged, and it leaves the wrong intact. Compared with political or legal practices such as pardoning, which suppresses the penal consequences of the offence, or amnesty, which consists in forgetting the wrong *ab initio*, the picture is quite different. Personal forgiveness is not a form of clemency that wipes out the offence or rules out penal responses.

I have suggested that the initial decision to forgive creates a new context, in which the process of forgiveness may thrive. One interesting difficulty in the analysis of forgiveness, which is also a clue to a philosophical solution, is that the term may stand for both the decision and the process. ‘I forgive you’ is an utterance that signals the decision to forgive and related conventions. However, one cannot reduce the whole story of forgiveness, the process of forgiving, or even the initial decision, to that speech act (contra Swinburne 1989, p. 85). Forgiveness as a decision as well as a process may remain silent, without the need for such an utterance. Whereas the performative ‘I promise’ is that in which my act of promising may consist, ‘I forgive you’ is not the whole of forgiveness, for two reasons: first, forgiveness involves more than the initial commitment; second, and most importantly, that commitment is usually signalled by means other than this utterance, notably by the behaviour of the forgiver (on ‘commissive forgiving’, as distinct from promising, see Pettigrove 2012, pp. 12–17). Moreover, sometimes (when one belatedly becomes aware of having already overcome one’s vindictive attitude) the decision to forgive is just reported either by a constative use of ‘I forgive you’ or by other means.

The utterance may have various functions, notably expressive (conveying one’s non-vindictive attitude), declarative (directly producing the state of affairs in which the wrongdoer is forgiven, which, were it

the only use of the formula, would be an argument for views that make forgiveness conditional on sovereign power), commissive (committing oneself to a specific course of action), but also assertive (stating a matter of fact). The two-stage account fits in with that flexibility, because the way in which the utterance may signal forgiveness varies according to its temporal distribution in the story. The commissive function operates at the beginning, and may be reiterated.

However, although it is interesting to approach forgiveness, or rather the initial step, from the angle of speech-act theory (Haber 1991), it cannot provide a full theory of forgiveness. Focusing exclusively on illocutionary aspects misses an important point. One consequence of the two-stage account (which does not apply to the divine or monarchical bestowal of grace, or the Roman Catholic absolution of sin) is that 'I forgive you' is not, strictly speaking, a performative utterance: forgiving is not something one does just in virtue of saying it (Novitz 1998, pp. 301–2). Perlocutionary effects on the forgiver as well as on the wrongdoer are crucial conditions of success.

In this sketch of personal forgiveness, I emphasize the process that follows the initial decision. However, there is also a prior process that forgiveness seems to require, a process whereby the forgiver and the wrongdoer arrive at a shared awareness of the gravity of the fault. Think of what instant forgiveness might look like. Imagine a situation in which someone seriously wrongs someone, so that the latter strongly resents the former, but almost immediately offers forgiveness. There is something weird about this. Forgiving instantly makes sense in some contexts, when no serious fault is involved (and thus no real forgiveness either), for instance, in the case of a breach of etiquette. To some, it is exquisitely polite to skate over another's *faux pas*, but forgiving does not mean ignoring the wrong. The kind of moving forward that forgiveness fosters is not that of oblivion, inattention or indifference. It takes some time for the wrongdoer, as well as for the wronged, to assess and assume the reality and seriousness of the wrong. Contrary to what accounts that recommend penance claim (see Swinburne 1989), the necessity of this preparatory process is not necessarily moral, but rather psychological and epistemic. Since the formal object of forgiveness is an apparent serious wrong, the thought of that object has to form, otherwise one would not know what one is doing when forgiving. Both parties need time to form the belief that the wrong is serious, so serious that it could justify revenge, and to acknowledge its reality.

Thus, the dimension of temporality is crucial to forgiveness, on the side of the forgiver as much as on the side of the wrongdoer. Some claim that forgiveness is conditional on the culprit's repentance. What I call the preparatory process does not necessarily consist in that. It is tempting to construe that time as a period during which wrongdoers go around in sackcloth and ashes, adopting the penitential attitude that proves they are not unworthy of forgiveness (see Swinburne 1989, pp. 81–4).<sup>2</sup> However, forgiveness is one possible response to a serious wrong, and requires a minimally shared recognition of the seriousness of the wrong, not a maximal condition of penance. That preliminary period of time, although it may result in the wrongdoer's repenting, is given to both parties so that they may arrive at a meaningful comprehension of the fault and a recognition of its reality, without which forgiving and being forgiven would lack sense and motivation. Time is required to take measure of the seriousness of the wrong. This is a normative point indeed, but epistemic rather than moral. One commonly accepted justification of declining to forgive is that it would be premature. If the wronged says to the wrongdoer 'it is too early', the latter may fill the waiting time by doing something, such as cultivating remorse or regret. Rather than being a necessary condition of forgiveness, penance seems to me a possible consequence of its temporal structure: forgiveness responds not to the event of the wrong, but to that wrong as an object of joint consciousness, which does not emerge instantly. However, there are particular situations in which penance seems to be a necessary condition: when the wrongdoer asks for forgiveness and claims to be worthy of it.

One more word about the condition of recognition. Let us consider an extreme case. Can one forgive a wrongdoer who does not simply deny that the wrong is a wrong (that is, who denies being guilty, for instance by claiming to be excused on some basis), but also does not acknowledge the bare facts (denies being the author of anything)? Without considering the moral question, there is obviously a psychological difficulty. In this particular situation, one can renounce revenge, but does that constitute forgiveness proper? Not if an expressive and communicative dimension is essential to

<sup>2</sup> For Swinburne, there is a prerequisite to forgiveness: atonement. It is not clear to me whether Swinburne considers atonement as a minimum or a maximum condition. Although he characterizes it as a 'small contribution' (Swinburne 1989, p. 81), in his view atonement includes repentance, apology, reparation, and (for serious wrongs) penance.

forgiveness, that is, if forgiveness ‘functions within the second-personal space of holding responsible’ (Darwall 2006, pp. 72–3) and is supposed to signal the relations between the wronged, the wrong and the wrongdoer. If the existence of the wrong and the involvement of the wrongdoer are deeply contested, if it is a case of a full denial of responsibility, that ‘social act of mind’—to hijack Thomas Reid’s phrase (Reid 2010, p. 330)—cannot work. The forgiver would need at least to anticipate minimal recognition on the wrongdoer’s part.

Although I stress this preliminary process before the decision to forgive and the process that is expected to ensue, the view I take of forgiveness is a two-stage, not a three-stage one. For during the preparatory period, forgiveness has not yet begun. The formation of a joint awareness of the wrong, vouching for its reality, might also lead to bitter resentment on one side, without any intention to forgive, and intense remorse on the other, without any claim to forgiveness, or to various combinations of reactive passions.

Let us get back to the equivocality of *forgiving*. It may stand metonymically for the initial act that opens it as well as the consequent process.<sup>3</sup> This is why there is no real contradiction if the same person, at different points in the same story, claims both to have forgiven and to be unable to forgive the same wrong. *Forgiving* may sometimes refer to the solemn decision and sometimes to the fallible process of transformation to which the decision commits the forgiver. On this view, forgiveness is not an interaction between two strong-willed and sovereign agents, nor between a free and generous forgiver and a weaker wrongdoer. It is a strategy (rather than a therapy<sup>4</sup>) that the forgiver employs to achieve the very difficult aim of transforming him or her attitudes and conduct, and of allowing the other, perhaps, to move forward. Thus, forgiveness is a tool for the weak agent.

Some claim that the transformational aim is about reconciliation. They do not mean that forgiving leads to friendship, which would be implausible, but that it aims at restoring or instituting relationships

<sup>3</sup> Charles Griswold observes that forgiveness may either refer to a ‘process’ or to an ‘end-state’ (‘fully achieved forgiveness’). According to his view, the process includes the moderation of resentment and a commitment to a non-resentful conduct (Griswold 2007, p. 42). The distinction I put forward is slightly but significantly different, since I situate the commitment to forgive before the ‘process’.

<sup>4</sup> The distinction between therapies as ways of healing and tactics or techniques as ways of coping is suggested by Schelling (1985, pp. 363–4).

between the forgiver and the wrongdoer. One objection to such a view is that forgiveness makes sense even in the absence of any relationships to restore and that it does not necessarily involve a desire to establish new relationships (Garrard and McNaughton 2003, pp. 45–7). An extreme version of the reconciliation theory would be the idea that forgiving replaces hate with love. This might be acceptable if ‘love’ is understood as practical rather than sentimental.<sup>5</sup> For there is an interesting ambiguity surrounding ‘love’ as well as the vindictive attitudes that love renounces. Does overcoming the vindictive stance mean renouncing anger and the desire to get one’s revenge, or just renouncing acting upon that desire or anger (on this Butlerian question, see Griswold 2007, pp. 38–43)? There is also a related question: if renouncing resentment may take some time, can one more immediately renounce cashing out one’s resentment in a vindictive course of action? The decision to forgive is a decision not to accept one’s desire to get even. That negative acceptance is practical and volitional. It does not directly suppress the vindictive feelings but does affect their practical consequences. Thus, it might be the case that someone has decided to forgive, acts upon that decision—that is, does not carry out revenge—and persistently feels hate, anger, or other negative emotions. This psychological fact is also consistent with the intentional structure of forgiveness. The formal object of forgiveness, a serious wrong, has to remain present as an object of joint awareness and memory throughout the whole process. To forgive is not to look away (on forgiveness as distinct from ‘condonation’ or ‘forgetting’, see Kolnai 1973). The forgiver resists a vindictive course of action and makes an effort to silence the inward advice of revenge, but preserves the memory of the wrong that has been done and may still have the feelings that go with it (Blustein 2010).

Personal forgiveness is supposed to operate upon revengeful desires. We need something more specific. Does the operation consist in diminishing, containing or suppressing desires? By ‘revengeful desires’, do we mean a set of emotional attitudes and feelings, or a conduct, a set of volitions and actions, or both? Forgiveness as an initial decision requires a situation in which we have hostile desires

<sup>5</sup> Glen Pettigrove interestingly stresses the connection between forgiveness and the volitional dimension of love (Pettigrove 2012, pp. 86–100. This shows at least that the final state of successful forgiveness may be achieved through volitional love, not that love is necessarily at work in forgiveness as a process.



against the one who has wronged us, and consists in deciding not to translate those desires into hostile actions against the wrongdoer. There are at least two different conceptions of the way in which one may alter desires that motivate retaliation. One conception involves directly controlling one's feelings; far from easy, except perhaps for a Stoic sage (who, however, being immune to vindictive sentiments, would not need to forgive, in the event of a wrong, which cannot affect a perfect Stoic). Another conception holds that although one cannot change one's desires at will, one can have a desire to act and not act according to that desire. I set aside the question of whether the ability to block or endorse one's desire should be viewed as evidence of free will. It suffices here to mention that the influence of other desires, not necessarily 'second-order' desires or volitions, may account for that ability.

In this dynamic account of forgiveness, the task of altering and perhaps extinguishing vindictive desires is left to the second stage, the subsequent process. The first stage, which suspends retaliation, has the further effect of opening up a process of affective transformation. Many times, the decision to forgive has the sole effect of triggering the slow reform of feelings, simply because the more immediate function of blocking vindictive action is running on empty, revenge being out of reach. By deciding to forgive, one creates a significantly different context, one in which one's attitudes and feelings are more likely to evolve. Many philosophers agree that forgiveness is a matter of overcoming one's hostile feelings (Garrard and McNaughton 2003, pp. 42–5). One advantage of the two-stage account is that it explains how forgiveness can perform that task. Were the operation of forgiveness confined to the nucleus of the decision to forgive, it would be an impossible performance, since one cannot efficaciously decide to change one's desires in the sense that one successfully decides to raise one's arm: deciding to forgive is more akin to deciding to become more sociable and friendly. The former, in normal conditions, is just a matter of willing, the latter is not.

Let us try to express this in one idiom of the philosophy of action. We ascribe to Jane, as a potential forgiver, several beliefs and desires, which constitute necessary conditions of forgiveness:

The belief B<sub>I</sub> that she has been wronged by Paul, and that he is thus responsible for a wrong that is serious enough to warrant,

not only blame, but also personal revenge. This is a belief that the decision to forgive and the subsequent process leave intact (Allais 2008), contrary to what happens in cases of amnesty, in which the wrong is forgotten.

The desire  $D_1$  to retaliate and get revenge, or at least feelings of anger and resentment. This desire, or rather this set of sombre affects and hostile motivations, is the material on which forgiveness operates, and cannot be controlled at will, directly, but only in an oblique way. This implies that someone exempt from  $D_1$  would not need to forgive. Of course, there are cases in which what we call forgiveness consists in just looking away and moving forward, in the absence of  $D_1$ ; this so-called forgiveness is oblivion, not memory of the wrong. Although the constellation of forgiveness is flexible enough to include many forms, this is closer to amnesty. There are also cases of hypothetical forgiveness: were  $D_1$  absent, a witness might nevertheless construe Jane's lenient conduct towards Paul as an instance of forgiveness.

The desire  $D_2$  to avoid the consequences of acting under the influence of  $D_1$ ; possibly the desire to overcome the feelings that accompany  $D_1$ , or the desire of having desires other than  $D_1$  and being the kind of person who does not take revenge; possibly other desires akin to practical love, such as giving Paul a second chance, etc.

More or less explicitly, the belief  $B_2$  that forgiving is what she should do in order to attain the object of  $D_2$ , and that deciding to forgive is a first step in that direction.

The combination of  $B_2$  and  $D_2$  leads Jane to engage in a course of action that has no immediate effect on  $D_1$ , but blocks the practical effects of the combination of  $B_1$  and  $D_1$ , and eventually may modify  $D_1$  itself.  $B_1$  and  $D_1$ , together with  $B_2$  and  $D_2$ , form the core of the input of forgiveness. What is fascinating is that all are still present in the output of the decision to forgive (as to desires, one still has  $D_1$ , since it cannot be suppressed by decree, and  $D_2$ , since the decision to forgive does not suffice to satisfy it); as to the final output of the process of forgiving, if it is successful,  $D_2$  is satisfied,  $B_2$  may subsist but has lost its practical relevance,  $D_1$  is altered, but  $B_1$  is left untouched. If the first decisional step does not modify the input (and it appears that it does not) the question is: what is the use of deciding to forgive?

My proposal is that deciding to forgive is equivalent to paying an entrance fee to the subsequent process. Although the output of the initial decision, in terms of beliefs and desires, appears to be identical to the input, there is some difference: B<sub>2</sub> and D<sub>2</sub> have exerted authority upon B<sub>1</sub> and D<sub>1</sub>, so that the former are validated and the latter rejected as principles of action. In other words, what is in the output that was not in the input is just the decision to forgive. The entrance fee consists of a non-vindictive behaviour. What is at stake is more than the practical rejection of D<sub>1</sub>: it is its revision. There are a variety of more or less successful outcomes, from blocking the consequences of D<sub>1</sub> to diminishing or even abandoning it.<sup>6</sup>

What is particular about the practice of forgiveness is its aim of changing one of the psychological states at its motivational core, that is, D<sub>1</sub>. To forgive is not to change one's desires or beliefs, at least not initially, nor to acquire new ones, at least not immediately. It is more akin to a pre-commitment, that is, a way of influencing one's future conduct—possibly against one's desires. There are several types of pre-commitment techniques, and a major divide between external constraints (Schelling 1992) and inward resolutions (Ainslie and Haslam 1992). Obviously, forgiveness as a pre-commitment belongs to the latter. The efficiency of internal resolutions or 'personal rules' is a controversial issue, but I think that the case of forgiveness pleads in favour of the thesis that with some luck they may be efficient.

One objection to the view that the decision to forgive is a pre-commitment, analogous to the payment of an entrance fee, is that there are cases of forgiveness (even successful ones) in situations where the decision to forgive is absent. My first answer is that the flexibility of the notion of forgiveness allows for forms in which the decision to forgive does not intervene, as well as forms in which the decision plays a major role in producing favourable circumstances. However, I propose that there is merely a difference in degree

<sup>6</sup> There is an interesting question about regretting forgiving. Jane may regret forgiving Paul for several reasons: because she no longer thinks that forgiving is the right thing to do (B<sub>2</sub> has vanished), or because she ceases to desire to be a forgiver (D<sub>2</sub> has vanished), or because she has acquired a new B<sub>1</sub>-type belief and a new D<sub>1</sub>-type desire, or because of a combination of these reasons, or for yet other reasons. Does regretting forgiving amount to reopening the file of revenge or creating a new file, such that regretting forgiving would be identical with ceasing to forgive? I think so. Regretting forgiving is quite different from regretting promising. The latter does not cancel the promise. One may fail to keep one's word, but one cannot fail to give it, when one gives it. On the contrary, forgiveness, as a process, is fallible and defeasible.

between the decision to forgive and the implicit acceptance of the premiss that one should not take revenge. Forgiveness may have its origins in an implicit and unconscious stance as well as in a solemn decision. Yet, it is doubtful that an implicit choice can be referred to as a pre-commitment in the specific form of a resolution. Adopting silently a particular course of action may nevertheless serve as a pre-commitment, as we shall see below.

An important qualification: the decision to forgive does not play only the role of a pre-commitment, because it also functions to communicate with the wrongdoer. A dynamic account should not underestimate the expressive dimension and the normative relations between the two parties. Specifically, on this view, the influence of the declaration on the wrongdoer is more direct than on the forgiver's own further transformation. The former is informed of the latter's commitment, which makes a difference, and can react to that. However, this explicit commitment does not give the wrongdoer an enforceable right to forgiveness, although it creates expectations on his or her part. Nor does the communicative element suppose the existence of a community of which both should be members, living together on good terms. It remains the case that, without the communicative element, which is conveyed either by speech acts or by discernible behaviour, forgiveness would amount to a technique of self-management, and would thus lose its relational dimension.

## II

*The Role and Influence of the Decision to Forgive.* Both the decision to forgive and the subsequent process may be described as consisting in renouncing revenge, but in different senses. The former is a commitment not to act in compliance with one's persistent vindictive desires, whereas during the latter the desires and negative emotions that go with them are altered and may be eventually extinguished. This account thus combines a practical and emotional understanding of 'renouncing revenge', whereas traditional accounts urge us to choose one or the other. Why not confine forgiveness to the first stage? Because, according to the language of forgiveness, what happens during the second stage matters to decide whether one did actually forgive.

My proposal is that the relation between the two stages is indirect. The decision to forgive does not directly cause the process of forgiveness. It creates circumstances that, along with some luck and other circumstances on the wrongdoer's side, may cause a transformation of feelings. What is new in this set of circumstances does not simply result from the decision to forgive, but also depends on favourable factors, such as the wrongdoer's attitude or other changes in the story or environment of the wronged, notably the removal of obstacles to any change of posture (for instance, the belief that one is an object of scorn), and all factors that may be summarized as 'luck'.<sup>7</sup> The process of forgiveness is encouraged by circumstances that are not wholly created by the decision.

However, forgiveness is not a blindfold adventure. It is a paradigmatic social practice or a quasi-institution in which roles are predefined. The forgiver draws on common knowledge about what kind of things might happen and what course of conduct is expected when one commits to forgive, although there is nothing automatic about this. To decide to forgive is quite different, in every respect, from deciding to 'Brexite': forgivers know where they are going, although they are never sure of getting there.

The requirements of forgiveness may be presented in term of the conditions necessary for forgiveness to occur:

On the forgiver's side, a necessary but not sufficient condition is the commitment to forgive, either as a speech act or as the practical acceptance of a non-vindictive line of conduct. This disjunctive necessary condition is not a sufficient condition, because the process of forgiveness depends on other factors.

On the wrongdoer's side, there are no sufficient conditions (otherwise the wrongdoer would have a right to be forgiven, once some conditions are fulfilled, which does not fit in with the common understanding of forgiveness); there is at least one necessary condition, that of considering the gravity of the wrong, which may consist in repentance or other ways of seeking forgiveness.

All these necessary but insufficient conditions, including luck and other necessary conditions, constitute as a whole an unnecessary and sufficient condition of forgiveness, to draw on John Mackie's

<sup>7</sup> On the 'vulnerability of forgiveness to luck', see Griswold (2007, pp. 130–3).

notorious analysis of the historians' use of 'cause' in terms of an 'insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result' (Mackie 1965, p. 245). The whole set of causes is unnecessary because the final effect, that is, the end state of forgiveness, might be produced by other sufficient sets of causes, such as love, divine intervention, mercy, chronic indulgence, the inability to hear what Jeffrie Murphy calls the 'legitimate claim' of vindictive passions (Murphy 2003, p. 117), and even a depressed sense of deserving any wrong done to us.

Now the question is whether the decision to forgive, as one of the insufficient but necessary causal ingredients, plays a special role, perhaps the main role, in bringing about the subsequent process and the final state. The question may sound bizarre. When I decide to go to the swimming pool next week, and I do so, my decision is obviously a relevant cause of my going to the swimming pool. This decision is also a reason to do so, in addition to the reasons for my decision (on decisions as 'reasons for performing the act decided upon', see Raz 1978, p. 138). It is at this point that the premiss of weak agency comes in. I describe as 'weak' those agents who have significant difficulty controlling their conduct over time, acting in accordance with their judgement as soon as they lose sight of well-considered reasons, sticking to their most solemn decisions. In fact, weak agents, in this sense, are not unable to forgive. Repeated evidence of their tendency not to keep resolutions may negatively affect their trustworthiness in the eyes of others, as well as their self-esteem and self-trust. However, it seems that it is not pointless for them to take decisions and plan their future conduct. The question in the case of the resolution to forgive is: how does it work?

According to the premiss of weak agency, the process of forgiveness is brought about by a set of favourable circumstances of which the commitment to forgive is a part. The process, properly speaking, is not the effect, but rather the consequence of that decision. Here I draw on Hart and Honoré's distinction between effect and consequence (Hart and Honoré 1959, p. 25). Event *N* may result from event *M* in at least two different ways. *N* may be the *effect* of *M*, *M* being a condition (simple or complex) sufficient to bring about *N*. *N* may be the *consequence* of *M*, *M* being not a condition sufficient to bring about *N*, but a necessary part of the complex condition that is responsible for the production of *N*. For instance, someone is fined for speeding. This is not the effect of the speeding in the way that the

melting of wax is an effect of heat, but is rather its consequence. It is an effect of the set of conditions that include the speeding along with the presence of police or a speed camera, the legal or administrative context and procedures, and so on.

Let us consider another example, which brings us closer to the point under discussion. In usual circumstances, my leaving a social event is an effect of my decision to leave, together with, say, a sentiment of boredom, the desire to go back home, or other motives. Even though the decision is closely connected with other determinants, including motives and reasons, it makes sense to say that my decision to leave is the most relevant cause of my leaving. Now, let us consider different circumstances. At 1 a.m., slumped on a sofa, sipping one more drink, soothed by ambient music, I am well aware that I ought to leave the party in order to be fresh enough for my early morning work. I decide that at 1:30 I will stand up and grab my jacket. At 1:50, I am still somewhere between the sofa and the front door, glass in hand, enjoying a very long series of goodbyes, unable to tear myself away, hypnotized by exquisite company. At 2:30, luckily, an abstemious guest offers to drive me back. In this case, the leaving is a consequence, not an effect, of the initial step. However, the decision was not useless, for it triggered a course of action that ended in my leaving. I might have stayed much longer without my 1:00 decision to leave the sofa, missing the opportunity of the late escort. Although my self-control was limited to the ability to decide to leave without extending further to the subsequent course of action, my precarious planning of my future conduct was lucky. My getting off the sofa was a way of putting myself in a position to act in spite of my desire to stay.

This may become more intelligible when we look at the distinction between mere *sine qua non* conditions and *sine qua non* conditions that are also causally relevant (Hart and Honoré 1959, p. 107). Enrolling as a student is a mere *sine qua non* condition of getting a degree. Although it is a necessary but insufficient condition of success, together with other conditions, it is not causally relevant as passing exams is. Buying a lottery ticket is a mere *sine qua non* condition for winning the lottery. Being the one whose number is drawn is another *sine qua non* condition, which is also insufficient per se (for other conditions are required, such as having the ticket in one's possession). Contrary to the initial purchase of the ticket, the conditions associated with the lottery draw are causally relevant.

The point is that, like logical requirements, statutory conditions have an analytic connection to the event. The fact that unless one is enrolled one cannot graduate, or that unless one is a participant one cannot win the lottery, is similar to Hart and Honoré's example of causally irrelevant *sine qua non* conditions in producing harm or a crime:

'If she had never married she would not have been a widow.' Such a remark would have a function as a *reminder* that it would not be correct to say that she was a widow if she had never been married. . . . Plainly, this is a condition *sine qua non* which it would be absurd to list among that infinite series of necessary conditions from which, according to modern juristic theory, we have to select the 'proximate cause'. (Hart and Honoré 1959, p. 108)

It is true that the problem with forgiveness, unlike the legal problem of identifying an offender, is not how to select, from a large set of necessary conditions, the one that is the proximate cause. In this matter, we know perfectly well who the culprit is, so to speak: the one who decided to forgive. In this context, as well as in legal contexts, especially criminal, the idea of 'relevant cause' has a normative dimension and points to agential responsibility. However, it also has a descriptive function and helps characterize the way in which an outcome is brought about. I draw on Hart and Honoré's distinctions between 'effect' and 'consequence' and between different types of necessary conditions only as descriptive tools, in order to account for the remote and unstable intrapersonal way in which the decision to forgive influences the future selves of forgivers who do not have firm control over their conduct over time.

My argument is that, under the premiss of weak agency, the decision to forgive, or any action taken in lieu of this decision, is a mere *sine qua non* condition. It removes obstacles from the path of forgiveness and sets up some favourable circumstances. It is also crucial to the expression and communication of the intention to forgive, and thus it may be causally relevant for the wrongdoer, as facilitating reform or atonement. However, in a context of weak agency (which admits of degree), it does not seem to be true that the decision to forgive per se plays a prominent role in contributing to successful forgiveness from the intrapersonal perspective. 'I forgive you' is not a magic formula that would open up the prospect of peace and automatically trigger the subsequent steps. Through the commitment to forgive, one both endorses the statutory role of a forgiver and enters



an experience that is not fully under one's control. A side consequence of the account I am proposing is that the practice of forgiving does not require high degrees of command and is not tailored to merciful heroes.

### III

*A Pascalian Conclusion.* To sum up, the same operation serves as a kind of self-nudging, by which one enrolls in the experience of forgiveness, and as a means of communicating with the wrongdoer. A decision to forgive provides the wronged with an incentive and a further reason to forgive, and creates a prospect of forgiveness for the wrongdoer. For non-weak agents, if they exist, the decision to forgive is the most relevant cause of successful forgiveness; it creates a normative expectation to which they respond, and it is the main reason why their attitudes and behaviour towards the culprit change. For weak agents (the rest of us), who have difficulty behaving in line with their important evaluations over time, the decision to forgive, although not as efficient, is not useless: it is a way of getting a foot in the door.

In Blaise Pascal's account of the 'wager', the libertine, like everyone, desires happiness and is convinced by the apologist that betting that God exists—that is, deciding to live on the premiss that God exists—is instrumental to happiness. The acquisition of faith is not an effect of the bet, but its possible consequence, since it involves other factors, mainly God's grace as the relevant cause. Faith cannot be obtained at will, nor be directly brought about by human means, and this is true not only of the libertine, but of all human agents, who in Pascal's view are constitutionally weak. Thus, the wager, which may consist in an implicit practical commitment to a Christian way of life as well as in an explicit decision, plays the qualifying role of an entrance fee, which, however, does not buy faith.

The dynamic view of forgiveness I propose draws on similar principles to Pascal's wager, although the former, contrary to the latter, is about how to change one's desires, not one's beliefs. Just as the wager is a useful technique for the libertine, not for the faithful, so people who need to forgive are those that are resentful, not those who are merciful. In both cases there is an outcome that is out of direct reach—that is, faith or successful forgiveness—and a decision (to behave as a Christian or to forgive), which is nothing more than

a technique for putting oneself in a different context, one from which the outcome may be achieved. In both cases, weak agents develop dispositions that do not fit into Aristotle's account of the robust virtues of the *σπουδαῖος* (the moral gentleman), and better correspond to his understanding of good habits at a very early stage of development, as suggested by Elster (1983, p. 53; see Aristotle, *Nicomachean Ethics* 1103B22–25). One important difference with the case of forgiveness is that the wager is inevitable and is an ongoing commitment: according to Pascal, everybody (even the faithful) lives on either of these premisses: that God exists or that God does not exist. Forgiveness, by contrast, responds to the event of particular wrongs, not to the human condition in general.

A connection between forgiveness and the wager may be made at another level too. It is not only that forgiveness shares some important formal psychological features with the wager, but also that it may be integrated, materially, as an ingredient of a higher-order wager. To some, who share with Thomas Reid the 'consciousness of the frailty of human nature' and the sense of having themselves 'often stood in need of forgiveness' (Reid 2010, p. 132), being able to forgive is also a way of becoming someone whose faults could be forgiven by God (Matthew 6:15) or by others. At this meta-level too, the relevant question is whether forgiving is a recipe for obtaining the expected outcome or just a matter of taking first steps in a transformation that only partly depends on us. Thus, theological controversies about grace and the mundane analysis of the dynamic of forgiveness shed light on one another.<sup>8</sup>

*Université Paris 1 Panthéon-Sorbonne*  
*Institut universitaire de France*  
UFR 10  
17 rue de la Sorbonne  
75005 Paris  
France  
*jaffro@univ-paris1.fr*

---

<sup>8</sup> I am grateful to Catherine Audard, Mikaël Cozik, Vanessa De Luca, Sarah Hutton, Eléonore Le Jallé, Pierre Livet, Emmanuel Picavet, David Wiggins, and especially the editor, Guy Longworth, for comments and questions. Some of the material of this paper was presented at the 'Lundis de la philosophie' seminar in Paris, Ecole Normale Supérieure, and benefited from Francis Wolff's comments.

## REFERENCES

- Ainslie, Georges, and Nick Haslam 1992: 'Self-Control'. In George Loewenstein and Jon Elster (eds.), *Choice Over Time*, pp. 177–209. New York: Russell Sage Foundation.
- Allais, Lucy 2008: 'Wiping the Slate Clean: The Heart of Forgiveness'. *Philosophy and Public Affairs*, 36(1), pp. 33–68.
- Blustein, Jeffrey 2010: 'Forgiveness, Commemoration, and Restorative Justice: The Role of Moral Emotions'. *Metaphilosophy*, 41(4), pp. 582–617.
- Darwall, Stephen 2006: *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Elster, Jon 1983: *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press; Paris: Editions de la Maison des Sciences de l'Homme.
- Garrard, Eve, and David McNaughton 2003: 'In Defence of Unconditional Forgiveness'. *Proceedings of the Aristotelian Society*, 103, pp. 39–60.
- 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., and Tony Honoré 1959: *Causation in the Law*. Oxford: Oxford University Press.
- Kolnai, Aurel 1973: 'Forgiveness'. *Proceedings of the Aristotelian Society*, 74, pp. 91–106.
- Mackie, J. L. 1965: 'Causes and Conditions'. *American Philosophical Quarterly*, 2(4), pp. 245–64.
- Murphy, Jeffrie G. 2003: *Getting Even: Forgiveness and Its Limits*. Oxford: Oxford University Press.
- Novitz, David 1998: 'Forgiveness and Self-Respect'. *Philosophy and Phenomenological Research*, 58(2), pp. 299–315.
- Pettigrove, Glen 2012: *Forgiveness and Love*. Oxford: Oxford University Press.
- Raz, Joseph 1978: 'Reasons for Action, Decisions and Norms'. In Joseph Raz (ed.), *Practical Reasoning*, pp. 128–43. Oxford: Oxford University Press. Originally published in *Mind*, 84 (1975), pp. 481–99.
- Reid, Thomas 2010: *Essays on the Active Powers of Man*. Edited by Knud Haakonssen and James A. Harris. Edinburgh: Edinburgh University Press.
- Schelling, Thomas C. 1985: 'Enforcing Rules on Oneself'. *Journal of Law, Economics, and Organization*, 1(2), pp. 357–74.
- 1992: 'Self-Command: A New Discipline'. In George Loewenstein and Jon Elster (eds.), *Choice Over Time*, pp. 167–76. New York: Russell Sage Foundation.
- Swinburne, Richard 1989: *Responsibility and Atonement*. Oxford: Clarendon Press.