Forgiveness and Weak Agency

LAURENT JAFFRO
FORGIVENESS AND WEAK AGENCY

LAURENT JAFFRO
UNIVERSITÉ PARIS I PANTHÉON-SORBONNE

MONDAY, 27 NOVEMBER 2017
17.30 - 19.15

THE WOBURN SUITE
SENATE HOUSE
UNIVERSITY OF LONDON
MALET STREET
LONDON WC1E 7HU
UNITED KINGDOM

This event is catered, free of charge & open to the general public

CONTACT
mail@aristoteliansociety.org.uk
www.aristoteliansociety.org.uk

© 2017 THE ARISTOTELIAN SOCIETY
Laurent Jaffro is professor of moral philosophy at Université Paris 1 Panthéon-Sorbonne since 2009 and a senior fellow of the Institut Universitaire de France since 2017. He is presently visiting professor at the University of Neuchâtel. He formerly taught philosophy at Blaise Pascal University and at the University of Nanterre. He is editor of Analyse et Philosophie, a series at Vrin. He has published on eighteenth-century British moral philosophy, especially on the third earl of Shaftesbury, and more recently on Thomas Reid. His current project focuses on ‘second-best ethics’, that is, ethics for agents who are chronically subject to weakness and have difficulties to align their conduct with their important values and commitments.

The following paper is a draft version that can only be cited or quoted with the author’s permission. The final paper will be published in Proceedings of the Aristotelian Society, Issue No. 1, Volume CXVIII (2018). Please visit the Society’s website for subscription information: aristoteliansomociety.org.uk.
THE ACCOUNT OF FORGIVENESS I put forward in what follows does not claim to be universal. It 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 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, etc. Second, accounts of forgiveness are often biased by moral considerations. For instance, authors who characterise personal forgiveness as being unconditional, not subject to some preliminary gesture or particular effort on the part of the offender, tend to see unconditionality as a demonstration of authentic generosity. Conversely, a view of forgiveness as necessarily correlated with penance is difficult to disentangle from the moral or religious argument that one ought not to forgive those who do not repent. However, what matters here is not what kind or degree of forgiveness a theory of forgiveness should value, but rather how to account for how forgiveness works. I leave aside considerations about the normative reasons 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 consider the forgiver as a weak agent, 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 forgiver. 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 as what Jon Elster calls by-products (1983, p. 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, or the decision to switch off the light leads to switching it off. 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.

In the first part of the paper, I sketch out the basic features of forgiveness that matter to my two-stage account of forgiveness. The second part tries to shed light on the relation between the two stages.

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, but the offender, who may be redeemed by forgiving. 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 (except when it is rudimental and remains just an alternative to revenge), where claimants expect from the institution something distinct from personal revenge: instead of trying themselves to get even with offenders, 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. On the contrary, in cases of personal forgiveness, the forgiver waives a vindictive course of action; since forgiveness is 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 and want him to be punished’, whereas it would be inconsistent to say ‘I forgive him and 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, 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 a performative 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 the act of promising consists, ‘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 may be signalled by other means, notably by the behaviour of the forgiver (on ‘commissive forgiving’, as distinct from promising, see Pettigrove 2012, pp. 12–7). 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. Although it is interesting to approach forgiveness, or rather the initial step, from the angle of speech acts theory (Haber 1991) it cannot provide a full theory of forgiveness.

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 there no other illocutionary uses, would be an argument for views that make forgiveness conditional on high degrees of autonomy), 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 (it may be reiterated), whereas the assertive is available at later moments, and the expressive may be employed at any stage.

In this sketch of personal forgiveness, I emphasise 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 forgivee 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 other’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 seriousness of the wrong. Contrary to what accounts that recommend penance claim (see Swinburne 1989), the necessity of this preparatory process is not moral but rather psychological. Since the formal object of forgiveness is a serious wrong, the thought of that object has to form, otherwise one would not know what one is doing when forgiving. One needs time to form the belief that the wrong is serious, so serious that it could justify revenge.

Thus, the dimension of temporality is crucial to forgiveness, on the side of the forgiver as much as on the side of the forgivee. 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 the wrongdoer goes around in sackcloth and ashes, adopting the penitential attitude that proves he/she is not unworthy of forgiveness (see Swinburne 1989, pp. 81–84). However, forgiveness is one possible response to a serious wrong, which 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, without

---

1 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 characterises it as a ‘small contribution’ (1989, p. 81), in his view atonement includes repentance, apology, reparation and (for serious wrongs) penance.
which forgiving and being forgiven would lack sense and motivation. Time is required to take measure of the seriousness of the wrong. 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), 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 forgiveness, that is, if forgiveness is supposed to signal the relations between the wronged, the wrong and the wrongdoer (on the ‘second-personal character’ of forgiveness, see Darwall 2006, pp. 72–3). If the existence of the wrong and the involvement of the wrongdoer are 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.

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 three-stage one. For during the preparatory period, forgiveness has not yet begun. The formation of a joint awareness of the wrong might also lead to bitter resentment on one side, and remorse on the other, 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. 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

2 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’.
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 forgivee. It is a strategy (rather than a therapy\(^3\)) that the forgiver employs to achieve the very difficult aim of transforming him/herself (his/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 between the forgiver and the forgivee. 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, p. 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.\(^4\) 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 [Joseph] 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 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 (see Kolnai 1973–4 on forgiveness as distinct from ‘condonation’ or ‘forgetting’). 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. The case of saints is different; they may know

---

\(^3\) The distinction between therapies as ways of healing and tactics or techniques as ways of coping is suggested by Schelling (1985, pp. 363-364).

\(^4\) 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.
full forgiveness, that is, a conversion of feelings as well as of practical attitudes.

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 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, does not need to forgive). 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 freewill. 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, p. 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 spend next summer in Greece. 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:

– As a necessary condition, the belief B1 that she has been wronged by Paul, the potential forgivee, and that he is 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 eventually forgotten.

– As a necessary condition, the desire D1 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 D1 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 D1; 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 D1 absent, a witness might nevertheless construe Jane’s lenient conduct towards Paul as an instance of forgiveness.

– As a necessary condition, the desire D2 to avoid the consequences of acting under the influence of D1; possibly, the desire to overcome the feelings that accompany D1, or having desires other than D1 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 B2 that forgiving is what she should do in order to attain the object of D2, and that deciding to forgive is a first step in that direction.

– Possibly, the desire D3 that Paul be punished by a third party. Whether this particular desire is satisfied or not has little impact on forgiveness. Conversely, it is doubtful that forgiveness operates on that desire. Indeed, a forgiving Jane may cease to desire that Paul be punished; however, in forgiving, she decides not to take her revenge, and this is distinct from her wishes about Paul being punished or not. Jane may both forgive Paul and be happy if he is penalised for his offence.

The combination of B2 and D2 leads Jane to engage in a course of action that has no immediate effect on D1, but blocks the practical effects of the combination of B1 and D1, and eventually may modify D1 itself. B1 and D1, together with B2 and D2, 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 D1, since it cannot be suppressed by decree, and D2, 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, D2 is satisfied, B2 may subsist but has lost its practical relevance, D1 is altered, but B1 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. Although the output of the initial decision, in terms of beliefs and desires, appears to be identical to the input, there is some difference: B2 and D2 have exerted authority upon B1 and D1, 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 — what one cannot expect without paying the fee — is more than the practical rejection of D1: it is its revision. There are a variety of more or less successful outcomes, from blocking the consequences of D1 to diminishing or even abandoning it.\(^5\)

What is particular about the practice of forgiveness is its aim of changing one of the psychological states at its motivational core, that is, D1. 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 cases 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. I do not claim to give a comprehensive

\(^5\) 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 (B2 has vanished), or because she ceases to desire to be a forgiver (D2 has vanished), or because she has acquired a new B1-type belief and a new D1-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 reversible.
account of all forms of forgiveness. However, I propose that there is merely a difference in degree between the decision to forgive and the implicit acceptance of the premise that one should not take revenge. Jonathan Cohen and others have contrasted acceptance and belief as respectively voluntary and involuntary attitudes. According to this kind of view, when I mechanically grab my umbrella before leaving home, I accept that it may be raining. Taking the umbrella may also be the object of an explicit decision (sometimes well-considered). In both cases, I act on the premise that it is, or may be, raining. Analogically, forgiveness may have its origins in an implicit and unconscious stance as well as in a solemn decision. However, 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 forgivee. 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 forgivee 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, the commitment does not give the forgivee an enforceable right to forgiveness. Nor does the communicative element suppose the existence of a community of which both are 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 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.

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 summarised as ‘luck’. The process of forgiveness is encouraged by circumstances that are not wholly created by the decision.

However, forgiveness is not a blindfolded 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 ‘Brexit’: 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 performative or as a 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 forgivee’s side, there are no sufficient conditions (otherwise the wrongdoer would have an enforceable 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 attitudes.

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 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 effect 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), a depressed sense of deserving any wrong done to us, etc.

Now the question is whether the decision to forgive, as one of the insufficient but necessary causal ingredients, plays a special role, perhaps

---

6 On the ‘vulnerability of forgiveness to luck’, see Griswold 2007, pp. 130–133.
the main role, in bringing about the subsequent process. The question may sound bizarre. When I decide to go to the swimming pool next week, 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 premise 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 premise 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 B may result from event A in at least two different ways. B may be the effect of A, A being a condition (simple or complex) sufficient to bring about A. B may be the consequence of A, A being not a condition sufficient to bring about B, but a necessary part of the complex condition that is responsible for the production of B. For instance, Isabelle, a reckless driver, is fined and sentenced for speeding. This is not the effect of her 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, etc.

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 1AM, 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, hypnotised 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 as causally relevant as passing exams, etc. 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 *per se* insufficient (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 159, 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 characterise 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 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 premise 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 forgivee, as facilitating reform or atonement. However, in cases 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. ‘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 enters an experience that is not fully under one’s control. Some have claimed that causation could be accounted for in terms of the manipulation of effects (Gasking 1955). To claim that A causes B amounts to asserting that a competent agent would be able to produce B by activating A. The first stage of forgiveness has nothing to do with such a formula. A side consequence of the account I am proposing is that the practice of forgiving is not restricted to only strong-willed people. It does not require high degrees of command and it 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 enrols 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, 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, 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 premise 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 as Pascal’s wager, although the former, contrary to the latter, is about how to change one’s desires, not one’s beliefs. The wager is a useful technique for the libertine, not for the faithful. Likewise, 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*, 1103b22–26). 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 premises: 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 who 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.

---

7 Some of the material of this paper was presented at the ‘Lundis de la philosophie’ seminar in Paris, École Normale Supérieure.
Université Paris 1 Panthéon-Sorbonne
UFR10
17 rue de la Sorbonne
75005 Paris, France
jaffro@univ-paris1.fr
REFERENCES


Reid, Thomas 2010: Essays on the Active Powers of Man, ed. K.


PRESIDENT: Helen Beebee (Manchester)

PRESIDENT-ELECT: Jo Wolff (UCL)

HONORARY DIRECTOR: Rory Madden (UCL)

EDITOR: Guy Longworth (Warwick)

LINES OF THOUGHT SERIES EDITOR: Scott Sturgeon (Oxford)

EXECUTIVE COMMITTEE: Clare Chambers (Cambridge) / Tim Crane (CEU) / James Harris (St Andrews) / Ulrike Heuer (Leeds) / Sacha Golob (KCL) / Elinor Mason (Edinburgh) / David Owens (KCL) / Jo Wolff (Oxford)

MANAGING EDITOR: Josephine Salverda (UCL)

ASSISTANT EDITOR: David Harris

DESIGNER: Mark Cortes Favis

ADMINISTRATOR: Nikhil Venkatesh (UCL)