**Forgiveness: From Conceptual Pluralism to Conceptual Ethics**

**Abstract**

Forgiveness theorists focus a good deal on explicating the content of what they take to be a shared folk concept of forgiveness. Our empirical research, however, suggests that there is a range of concepts of forgiveness present in the population, and therefore that we should be folk conceptual pluralists about forgiveness. We suggest two possible responses on the part of forgiveness theorists: (1) to deny folk conceptual pluralism by arguing that forgiveness is a functional concept and (2) to accept folk conceptual pluralism and focus on a revisionary conceptual ethics project.

**1. Introduction**

Whatever exactly forgiveness is (or should be), it is an important inter-personal and social phenomenon. Its presence, or absence, has significant psychological, social, and potentially political, consequences. So, seeking to know more about that phenomenon, and to know in what circumstances we find instances of forgiveness—engaging with the descriptive project of articulating the content of the concept deployed—is surely a worthwhile endeavour, and is something forgiveness theorists have keenly pursued.

Having catalogued disagreement between theorists, along a number of axes, regarding the content of the concept of forgiveness (§2) we present our recent exploratory empirical findings about the content of the folk concept (§3). We argue that our results provide evidence in favour of folk conceptual pluralism about forgiveness—the view that there are multiple distinct concepts used by ordinary people under the banner ‘forgiveness’. This presents a challenge for those philosophers offering accounts of forgiveness who aim to delineate a single, unified concept. We consider two ways forgiveness theorists might respond to this data (§4). They might argue that forgiveness is a *functional* concept, and that the disagreement we report is really only disagreement about what plays the relevant functional role(s). Or they might accept folk conceptual pluralism and turn their attention towards the project of determining which of these concepts we *ought* to deploy. With this in mind, we gesture towards some ways in which our data can be used to evaluate the practical prospects of projects in conceptual ethics.

**2. Extant Accounts of Forgiveness**

In everyday life we encounter many disagreements about when and why people should forgive. Christians, Hindus, Buddhists, and indeed therapeutic psychologists typically recommend that we forgive those who wrong us, even if they are unrepentant. Yet, at least some of the time, this kind of unconditional forgiveness seems unjust, premature, and possibly dangerous. Most philosophers who write about forgiveness hope to answer this normative question: When (and why) should we forgive? Obviously, though, we cannot do this without first answering a descriptive question: What is forgiveness? As we shall see, philosophers have responded to this descriptive question by offering a plethora of competing accounts of forgiveness.

Critics might claim that these philosophers are simply inventing definitions or are merely stipulating what they will mean by the word ‘forgiveness’. But this is not what these philosophers take themselves to be doing. Rather, they are aiming to give an accurate definition of *forgiveness*, the very thing that preachers and therapists have been talking about. Philosophical accounts of forgiveness are designed to latch onto and delineate an existing folk concept, not to invent a new concept, nor to articulate a range of distinct concepts that go by the name ‘forgiveness’.[[1]](#footnote-1)

These philosophers typically assume that we can identify some necessary and/or some sufficient conditions for forgiveness, and that we can test which conditions are necessary and which sufficient by consulting our shared intuitions regarding concrete examples, case studies, and realistic thought experiments.[[2]](#footnote-2) Some accounts of forgiveness are explicitly framed in terms of jointly necessary and sufficient conditions.[[3]](#footnote-3) Many philosophers who shy away from offering a full set of necessary and sufficient conditions nonetheless engage in debates as to whether a particular disputed condition is necessary or sufficient.[[4]](#footnote-4) These philosophers may admit that folk talk about forgiveness needs to be tidied up and made consistent, but at a minimum they aim to respect the central folk practices in their own fully developed accounts.[[5]](#footnote-5)

This methodology makes sense if we assume that there is a unified folk concept of forgiveness that is the target of these philosophical accounts. If there is a single shared folk concept of forgiveness, then we should be able to evaluate philosophical accounts of forgiveness according to how well they fit with the folk concept. It is natural to interpret many recent philosophical disagreements regarding forgiveness as instances of this methodology. For example, Griswold (2007) rejects accounts of forgiveness which deny the folk intuition that forgiveness is incompatible with continued anger. In turn, Allais (2008) rejects Griswold’s account of forgiveness because it fails to respect the folk intuition that it is possible to forgive the dead. These philosophers take themselves to be offering competing accounts of the same thing, not merely offering different definitions of distinct things. Each theorist thinks that his or her own definition of forgiveness is superior to those proposed by their opponents because it is descriptively more accurate: it gets at what forgiveness really is.

There are a few cases in which philosophers have expressed doubt as to whether there really is a single target at which all of these accounts are aimed. Pettigrove and Allais acknowledge that there might be several distinct concepts that are in common use amongst the folk, or at least several overlapping conceptions of forgiveness that are employed by different cultural groups (e.g. Allais 2008; Pettigrove 2012). Pettigrove (2012) suggests that if folk intuitions are too varied to help us decide between competing accounts, then we should develop an account of the kind of forgiveness that we might hope to find (i.e. a concept of forgiveness that it is overall maximising to have deployed in a society). Boleyn-Fitzgerald (2002) goes further than this, arguing that we should adopt his preferred conception of forgiveness not because it is the closest fit with folk intuitions, but because it will help us to let go of our anger.

These gestures towards normative or pragmatic criteria are tantalising. Yet if it turns out that there is a unified folk concept of forgiveness, philosophers who diverge from this concept on normative or pragmatic grounds will have to answer to the charge of wishful thinking. We want to find out what forgiveness really *is*, not merely imagine some nice things that we could choose to call ‘forgiveness’. Given the state of play in the philosophical literature, it would be very useful to figure out whether there is a unified folk concept of forgiveness, and thus a single descriptive target for our philosophical theorising.

Before moving on to our empirical research into the folk concept, we should survey recent philosophical debates about forgiveness, and identify the common ground and the points of disagreement. There is broad agreement amongst philosophers that forgiveness is something that can occur after a victim judges that she has been wronged, and that forgiveness is one way in which the victim can ‘get over’ the perceived wrong. There is almost universal agreement that forgiveness is wholly distinct from various other ways of getting over having been wronged, such as excusing the wrongdoer, merely forgetting that the wrong took place, or coming to judge that the seemingly wrong action was in fact justified. What distinguishes forgiving from excusing, forgetting, and justifying, is the following necessary condition: after she has forgiven the wrongdoer, the victim still judges that the action was wrong and that the perpetrator was culpable. (Kolnai 1973-1974; Murphy 1988; Hieronymi 2001; Griswold 2007; Allais 2008; Pettigrove 2012). Let us call these the *culpability judgments*. The persistence of the culpability judgments is almost universally agreed to be necessary for forgiveness (see Boleyn-Fitzgerald 2002 for a rare conflicting view). Disagreement arises amongst philosophers as they attempt to spell out further necessary and/or sufficient conditions for forgiveness.

Many accounts suppose that a victim counts as having forgiven a perpetrator only if that victim has let go of certain negative emotions or attitudes that she initially felt towards the perpetrator as a response to the wrong action. These might include comparatively simple episodic emotions, such as anger or sadness (Boleyn-Fitzgerald 2002). Yet most forgiveness theorists claim that the relevant category is not that of negative emotion, but negative reactive attitudes. Reactive attitudes, following Strawson (1962), are affective dispositions that we direct towards agents in light of the quality of will that they demonstrate in their actions. For example, gratitude is a positive reactive attitude that we direct towards agents for their goodwill and their helpful actions. When it comes to wrongdoing and forgiveness, the key reactive attitude is taken to be resentment, a moralised anger that we feel towards those who have wronged us *because* they have wronged us (Hieronymi 2001; Allais 2008; Pettigrove 2012). Part of what it is to feel these reactive attitudes, as opposed to more basic emotions, is to judge that the agent is morally responsible for the action in question. Thus, it is possible to feel simple negative emotions in response to having been hurt without resenting the person who hurt you: you might be angry and sad that the incident occurred, but think that it was not really the perpetrator’s fault.

Some theorists hold that letting go of negative emotions is necessary for forgiveness (Boleyn-Fitzgerald 2002; Novitz 1998). Others claim that letting go of hostile negative reactive attitudes, especially resentment and contempt, is necessary for forgiveness (Murphy 1988, 2001; Hieronymi 2001; Pettigrove 2012; Allais 2013; Garrard and McNaughton 2002; Griswold 2007; Holmgren 1993; Richards 1988). In some cases, this condition is couched explicitly in terms of the *letting go* of these negative reactive attitudes, and thus if the victim never felt resentment to begin with, it is impossible for her to forgive (Griswold 2007; Murphy 1988; Novitz 1998). Others hold that it is the stable absence of negative reactive attitudes that is required for forgiveness, and hence that victims can forgive even if they never felt resentment in the first place (Pettigrove 2012). In contrast, some philosophers claims that forgiveness is compatible with the victim continuing to resent the perpetrator, so long as this continued resentment is moderated to an appropriate level (Butler 1726/2002), or is accompanied by the victim’s resolution not to let it affect the way in which she treats the perpetrator (Zaibert 2009A, 2009B; Warmke 2016).

At least one philosopher holds that the letting go of (relevant) negative emotions, for whatever reason, is *sufficient* for forgiveness (Boleyn-Fitzgerald 2002). On this view, forgiveness is equivalent to simply getting over it. We have seen that the overwhelming majority disagree, claiming that forgiveness is more complex than this. The dominant view is that after forgiveness has taken place, the victim must maintain her culpability judgments. This is the condition that ensures that there is no overlap between forgiving and mere forgetting, between forgiving and excusing, and between forgiving and justifying. Richards (1988) holds that maintaining the culpability judgments while letting go of both negative emotions and resentment is both necessary and sufficient for forgiveness.

Many other philosophers argue that more is required. Some have argued that forgiveness occurs only when the victim lets go of negative emotions and reactive attitudes via the right sort of process. For instance, it has been suggested that it is necessary that the victim let go of negative emotions and reactive attitudes by working through them in a way that allows the victim to recover self-esteem and cultivate good will towards the perpetrator (Holmgren 2012; Hampton 1988; Garrard and McNaughton 2002; Pettigrove 2004) or by fostering compassion for the perpetrator and gaining an empathetic understanding of his point of view (Novitz 1998).

Still others argue that forgiveness occurs only when the victim lets go of negative emotions and reactive attitudes *for the right reasons*. On some such views, forgiveness requires the victim to cease resenting upon correctly perceiving that the perpetrator is no longer a fit object of resentment. At a minimum, this will occur when the perpetrator has distanced himself from the wrong action via remorse and apology (Hieronymi 2001; Couto 2016), but some believe that it also requires the perpetrator has seen the wrong from the victim’s point of view, has properly compensated the victim, and has undergone moral reform such as to now be the sort of person who would not repeat such an action (Griswold 2007:50). Other theorists disagree, holding that so-called unconditional forgiveness can occur regardless of whether the perpetrator is repentant or apologetic (Pettigrove 2012; Allais 2013; Govier and Verwoerd 2002; Holmgren 1993, 2012; Garrard and McNaughton 2002; Zaibert 2009A).

Some philosophers claim that forgiveness has taken place only when it has been explicitly communicated by the victim to the perpetrator (Griswold 2007; Warmke 2016).[[6]](#footnote-6) Others believe that forgiveness can occur privately, without communication, and that we can forgive the dead, who are beyond our communicative reach (Allais 2008; Pettigrove 2012).

A final point of disagreement concerns whether, after forgiveness, victims must treat perpetrators differently. Some philosophers suggest that after forgiveness the victim cannot continue to punish the perpetrator (Hieronymi 2001; Zaibert 2009A, 2012). Zaibert (2009A) thinks that intentional refusal to punish, combined with maintenance of the culpability judgments, is both sufficient and necessary for forgiveness. Others hold that forgiveness has taken place only if the victim and the perpetrator are no longer in conflict over the proper response to the wrong. On this view, forgiveness *is* compatible with continued punishment by the victim, but only when the perpetrator accepts that punishment (Russell 2016). Others maintain that victims often have forgiven perpetrators while continuing to inflict ongoing punishment on them, whether the perpetrators accept that punishment or not (Warmke 2011, 2013; Haber 1991; Griswold 2007:32-33; Murphy 2003:101; Garrard and McNaughton 2002; Novitz 1998; Pettigrove 2012; Allais 2013).

**3. Our Study**

In order to make some headway resolving these disputes between forgiveness theorists we designed and conducted an exploratory policy-capturing study to investigate the ways in which the folk weight different factors in determining whether a scenario counts as one in which their concept of forgiveness is satisfied.[[7]](#footnote-7) Policy-capturing studies use regression techniques to determine how individuals weight a range of factors—known as *cues*—when making a decision: in this case, deciding whether or not a scenario is one in which forgiveness has taken place. Participants in a policy-capturing study are presented with a set of profiles—in our case a series of vignettes—in which the cues—the various factors forgiveness theorists are interested in—are varied systematically. Participants are then asked to assess the profiles: in our case to assess whether or not forgiveness has taken place. The data are then analysed via multiple separate regressions on each participant’s data. The standardised regression coefficients (beta weights), which we report, provide information about the relative importance of the cues to the participant, in making their decision about whether forgiveness has occurred.

Importantly then, a policy-capturing analysis does not aim to determine which cues are necessary, and which sufficient, for a participant to judge that forgiveness has occurred: instead, it aims to determine which cues are more, or less, important. So, we cannot neatly map these results onto claims about the necessity and sufficiency of these cues. Such an analysis will not tell us whether cues that receive a relatively high weighting are singly necessary, or sufficient, for forgiveness, or are jointly so. Nevertheless, it is safe to assume that cues that have *very* low weighting will not be sufficient for forgiveness, and likely will not be necessary either. So, a policy-capturing analysis may go some way towards answering questions about the necessity or sufficiency of certain cues. Further, we can determine whether the cues that forgiveness theorists have identified as important, are *in fact* cues that are important to the folk when they make decisions about whether forgiveness has occurred. Finally, and most importantly for our purposes, data gleaned from a policy-capturing analysis provides information about the variability in the population tested, regarding the weighting of the various cues. Hence it provides data that speaks to the issue of whether it is likely that there is a single shared concept of forgiveness amongst the population tested.

**3.1 Methodology**

Participants were presented with vignettes and asked to report whether a scenario was one in which forgiveness had taken place. This study design differs from a number of recent studies in psychology on the concept of forgiveness, in which participants were asked to list features associated with forgiveness (Friesen and Fletcher 2007) or asked how they would describe forgiveness to someone unfamiliar with the concept (Kearns and Fincham 2004). Such methodologies are, arguably, not well placed to determine the content of concepts, since individuals are often not able to articulate said content (see Knobe and Nichols 2008). Even putting experimental philosophy aside, many philosophers think that it is often impossible to give a neat conceptual analysis of our concepts. Indeed, even philosophers who think we have *a priori* access to the content of our concepts often think this is so.[[8]](#footnote-8) This is why we asked participants to *use* their concept in making judgements about cases.

The vignettes with which participants were presented varied systematically in which cues were present. It was not possible to consider all the factors theorists have put forward as important to forgiveness, so we focussed on six cues that often appear in the literature. We did not test the widely shared assumption that the persistence of the culpability judgments is necessary for forgiveness. In all our tested scenarios, the victim still judged that she had been wronged and that the perpetrator was culpable. Participants were asked to respond to each vignette on a 7-point Likert scale with a value of 1 indicating that they were completely sure forgiveness had *not* taken place, and a value of 7 indicating that they were completely sure forgiveness had taken place. We also asked whether or not a declaration of forgiveness must be in order for one party to count as forgiving the other.

*Table 1. Forgiveness cues*

|  |  |
| --- | --- |
| **Cue** | **Description** |
| **~nE[[9]](#footnote-9)** | Negative Emotion: The victim lets go of the negative emotion of anger, which the victim felt as the result of the perpetrator’s wrongdoing. |
| **~rA[[10]](#footnote-10)** | Reactive Attitude: The victim lets go of the negative reactive attitude of resentment towards the perpetrator, which the victim felt as a result of the perpetrator’s wrongdoing. |
| **cG** | Cognition: The victim works through negative emotions and reactive attitudes by correctly identifying the wrong done to him/herself, and identifying the source of negative emotions and reactive attitudes. |
| **~pT** | Punishment: The victim does not intend to have the perpetrator punished. |
| **bR** | Justified belief about remorse: The victim justifiably believes the perpetrator is remorseful. |
| **R** | Remorse: The perpetrator is remorseful. |

We tested every possible combination of those cues. Combinations of cues consistent with contemporary philosophical accounts of forgiveness include the following: ~nE[[11]](#footnote-11), ~rA[[12]](#footnote-12), cG[[13]](#footnote-13), ~pT[[14]](#footnote-14), bR[[15]](#footnote-15), R[[16]](#footnote-16), ~nE & ~rA[[17]](#footnote-17), ~nE & ~rA & cG[[18]](#footnote-18), ~nE & ~rA & cG & bR & R[[19]](#footnote-19), ~rA & cG & bR & R[[20]](#footnote-20) and ~nE & ~rA & cG & ~pT & bR & R[[21]](#footnote-21).

The concept of forgiveness was probed across two participant samples; each tested under either high or low valence conditions. Low valence scenarios involved stealing a hat, while high valence scenarios involved murdering the entire population of a village (for an example of each see Table 2, below). Those vignettes are available to view on our project page on the Open Science Framework (<https://osf.io/x34x5/>).[[22]](#footnote-22)

*Table 2. An example of a low valence and a high valence forgiveness scenario.*

|  |
| --- |
| Low Valence Forgiveness Scenario:Jane and Phoebe know one another, but are not close friends. Jane has a hat that Phoebe really likes. One day, Jane puts the hat down and Phoebe picks it up. Jane asks if anyone has seen her hat, and Phoebe responds that she has not. Phoebe takes the hat home. One day Jane is walking down the street and sees Phoebe wearing her hat. Jane confronts Phoebe. Phoebe admits that she picked up Jane's hat and took it home with her – that she stole Jane’s hat.Jane tells her mother that Phoebe stole the hat. Jane's mother suggests that Jane tell the schoolteacher about the hat-stealing, so Phoebe can be punished for her action - as stealing is wrong. Jane decides that doing so would be a good idea. She feels that what Phoebe did was wrong, and Phoebe ought not to have done it. If Jane were to tell the schoolteacher, Phoebe would get detention for a few days. At school the next week, Jane sees Phoebe. Jane is still angry. Moreover, she still resents Phoebe for stealing the hat. Upon reflection, Jane realises that this is because Phoebe stole her hat, and that stealing is wrong, and Phoebe ought not to have done it. As a result, Jane intends to have Phoebe punished in detention by telling the school-teacher.Jane sees Phoebe the next day, and notices that she is being very considerate, and appears remorseful. Jane comes to believe that Phoebe is sorry for stealing the hat, and regrets her actions. As it turns out, Jane is right. At the same time, Phoebe has been thinking about her action. Phoebe realises that she regrets taking Jane's hat. It was upsetting to Jane. She wishes she had not done it, and if she could go back and change it, she would. Phoebe promises to herself that in future she will not make the same mistake and take other people’s things. |
| High Valence Forgiveness Scenario:Jane and Phoebe know one another, but are not close friends. Jane grew up in a remote little village that she really likes. Jane and Phoebe work at a bio-tech company and both have access to some deadly pathogens. One day Jane puts down a deadly pathogen. Phoebe picks up the pathogen, because she knows that if she releases it in the village where Jane grew up, she will kill all its inhabitants and this thought gives her great pleasure. Phoebe takes the pathogen to the village and releases it, where it spreads throughout the village, killing everyone who lives there. Shortly after, Jane is walking into her home village and sees that everyone is dead. Jane confronts Phoebe. Phoebe admits that she picked up the deadly pathogen and released it in the village – that she murdered everyone.The CEO of the bio-tech company emails its employees suggesting whoever knows who released the pathogen should tell the authorities so the perpetrator can be punished for their action – as murder is wrong. Jane decides that doing so would be a good idea. She feels that what Phoebe did was wrong, and Phoebe ought not to have done it. If Jane were to tell the authorities, Phoebe would get jailed for life.At the bio-tech company the next week, Jane sees Phoebe. Jane is still angry. Moreover, she still resents Phoebe for murdering the villagers. Upon reflection, Jane realises that this is because Phoebe murdered the villagers where she grew up, and that murder is wrong, and Phoebe ought not to have done it. As a result, Jane intends to have Phoebe punished by telling the authorities and having her sent to jail. Jane sees Phoebe the next day, and notices that she is being very considerate, and appears remorseful. Jane comes to believe that Phoebe is sorry for murdering the villagers, and regrets her actions. As it turns out, Jane is right. At the same time, Phoebe has been thinking about her action. Phoebe realises that she regrets murdering the villagers where Jane grew up. It was upsetting to Jane. She wishes she had not done it, and if she could go back and change it, she would. Phoebe promises to herself that in future she will not make the same mistake and release the deadly pathogen in other villages. |

Scenarios were generated in a systematic manner by adding a pre-worded exemplar of each of six forgiveness cues in a ‘present’ or ‘absent’ form to a base scenario. Thus, *each cue* is mentioned in *each scenario*, in one of two forms. No scenario is simply ‘silent’ on a cue. So, for example, the cue of negative emotions could take the form nE, ‘Jane is still angry’, or ~nE, ‘Jane is no longer angry’. Each scenario includes either nE or ~nE. Similarly, each scenario includes either rA, ‘Jane still resents Phoebe for […]’, or ~rA, ‘Jane no longer resents Phoebe for […]’. We took anger to be a paradigmatically negative emotion, and resentment to be a paradigmatically negative reactive attitude. Hence, they are the only ones evaluated in this study.

Continuing on, the victim could have the cognitive ability to work through their negative emotions (cG), ‘Upon reflection, Jane realises that this is because Phoebe […], and that […] is wrong, and Phoebe ought not to have done it’, or not (~cG), ‘Jane is not sure why she is angry and resentful; she is unable to think clearly about what happened at all’. The victim could have the intention to withhold punishment (~pT), ‘Jane intends not to have Phoebe punished’ or could intend to have it administered (pT), ‘Jane intends to have Phoebe punished’. The victim could have a justified belief that the perpetrator is remorseful (bR), ‘Jane sees Phoebe the next day, and notices that she is being very considerate, and appears remorseful. Jane comes to believe that Phoebe is sorry for […], and regrets her actions’, or a belief that the perpetrator is not remorseful (~bR), ‘Jane sees Phoebe the next day, and notices that she is being very considerate, and appears remorseful. Jane thinks this is probably an act, and comes to believe that Phoebe is not sorry for […], and does not regret her actions’.

Finally, the perpetrator, as a matter of fact, could be remorseful (R), ‘At the same time, Phoebe has been thinking about her action. Phoebe realises that she regrets […]. It was upsetting to Jane. She wishes she had not done it, and if she could go back and change it, she would. Phoebe promises to herself that in future she will not make the same mistake and […]’, or not (~R), ‘At the same time, Phoebe has been thinking about her action. Phoebe realises that she does not regret […]. She is glad that she […]’.

3.1.2. *Participants*

Participants were recruited through advertisement on Amazon’s Mechanical Turk. Although no strict demographic exclusion criteria were enforced, all participants were over the age of 18—the age required to open a Mechanical Turk account; and all participants resided in the USA—the only country where monetary compensation can be paid. Ethical approval for this study was obtained from the BLINDED Human Research Ethics committee. Consent was obtained from participants on return of a completed survey. All participants were naïve to the study’s hypotheses.

Since there are no previous studies examining folk intuitions about the cues that philosophers have identified as important for forgiveness, we follow Green (1991) who recommends that for a medium effect size (standardised beta coefficient = .2), with 80% power, the sample size should be ≥104 + *p*, where *p* is the number of independent predictors. As we are using a total of 6 independent predictors this results in minimum sample size of 110. As our study includes both a low and high valence condition, the minimum overall required sample size is 220. A total of 243 fluent English speakers participated in the study in return for monetary compensation. Participants in the high valence condition saw, in randomised order, all the high valence vignettes, and participants in the low valence condition saw, in randomised order, all the low valence vignettes.

Data integrity was achieved through attentional checks. Data from 23 participants were excluded from the high valence survey and 22 participants from the low valence survey for failing to correctly answer an attentional check question. In addition, 2 participants were excluded from the high valence survey and 2 participants from the low valence survey for failing to answer all the questions. This gave us a final sample size of 101 for the high valence survey and 93 for the low valence survey. There were no statistical differences between survey samples for gender, age, and religious belief (see Table 3).

*Table 3.* *Participant* *demographics by survey sample.*

|  |  |  |
| --- | --- | --- |
| **Demographic** | **Low Valence** | **High Valence** |
| Gender (M/F) | 53/40 | 48/53 |
| Mean Age (S.D.) | 38.99 (10.91) | 38.48 (11.27) |
| Religious Belief (Y/N) [[23]](#footnote-23) | 48/45 | 52/49 |

**3.2 Results**

There are four key results that emerge from our experiments. First, our results showed large variation *between participants* regarding the weightings of the forgiveness cues. Second, there was variation in the *mean* weighting given to each cue (i.e. not all cues were weighted equally). The largest weightings were given to ~rA and bR, followed by ~nE and ~pT, and then cG and R. Third, there was variation in the *direction* of the mean weightings (i.e. not all cues were positively associated with forgiveness judgements). In particular, cue ~pT was weighted negatively, while all other cues were weighted positively. Finally, we found no association between the magnitude of forgiveness cue weightings or their direction, and valence.

In this study, we were primarily interested in exploring the standardised beta coefficients of each participant for each of the forgiveness cues being examined, and their association with valence. (Recall that standardised beta coefficients provide information about the relative importance participants ascribe to each cue when making their forgiveness judgments.) To do this we calculated two sets of within-subjects regressions: one set for participants in the low valence condition, and the other set for participants in the high valence condition. In both cases, we calculated the regression equation for each participant, with forgiveness judgments serving as the criterion variable and the forgiveness cues of ~nE, ~rA, cG, ~pT, bR, and R serving as predictors. This resulted in the creation of a total of 93 regression equations for the low valence condition, and 101 regression equations for the high valence condition. It was not possible to include interaction terms because of the low power afforded by our sample size.

First, let’s consider the low valence condition. As predicted, there was large variability in the extent to which the linear combination of forgiveness cues predicted each participant’s forgiveness judgments. The R2 for each participant ranged from .04 to .98, with a mean of .62 (*S.D.* = .22).[[24]](#footnote-24) The percentage of standardised beta coefficients that were statistically significant for each cue across participants was as follows: ~nE (50%), ~rA (60.2%), cG (17.2%), ~pT (50.5%), bR (51.6%), and R (14%). In other words, for 50% of our participants, forgiveness judgements were significantly associated with the presence of ~nE, and so on for the other factors. The mean and standard deviation of the standardised beta coefficients in the low valence condition can be found in Table 4 below.

Second, the high valence condition. Once again, as predicted, there was large variability in the extent to which the linear combination of forgiveness cues were associated with each participant’s forgiveness judgments. The R2 for each participant ranged from .035 to .946, with a mean of .55 (*S.D.* = .23). The percentage of standardised beta coefficients that were statistically significant for each cue across participants was as follows: ~nE (54.5%), ~rA (65.3%), cG (17.8%), ~pT (50.5%), bR (46.5%), and R (11.9%). In other words, for 54.5% of our participants, forgiveness judgements were significantly associated with the presence of ~nE, and so on for the other factors. Once again, the mean and standard deviation of the standardised beta weights in the high valence condition can be found in Table 4 below.

Setting aside consideration of valence for the moment, only the forgiveness cue of giving up resentment was found to be important to the majority of participants (62.9%). Giving up anger (52.6%), believing the perpetrator has repented (49%), and intending not to punish (50.5%) were important to only half of the participants. Finally, the victim working through their negative emotion and reactive attitude and identifying their source (17.5%) and the perpetrator being remorseful (12.9%) were important to only a minority of participants.

*Table 4. Mean standardised beta coefficients and standard deviations for each cue for each valence.*

|  |  |
| --- | --- |
|  | **Cue** |
| **Valence** | **~nE** | **~rA** | **cG** | **~pT** | **bR** | **R** |
| Low (n = 93) | .22 (.24) | .30 (.27) | .06 (.12) | -.21 (.22) | .32 (.33) | .08 (.18) |
| High (n = 101) | .20 (.23) | .31 (.28) | .06 (.10) | -.23 (.23) | .26 (.28) | .06 (.15) |

From Table 4 we can see that overall, participants weight the forgiveness cues ~rA and bR the highest, followed by ~nE and ~pT. Importantly though, and contrary to Zaibert (2009A), the forgiveness cue ~pT, overall, is weighted negatively: that the victim does not intend to punish the perpetrator weighs *against* the judgment that forgiveness has taken place. Finally, participants weight cues cG and R the smallest. Interestingly, there was no noticeable effect of valence, which we report in table 5.

In addition, we see a large amount of variability present in the overall results (which should not be surprising given the large amount of variability in the significance of forgiveness cues between participants; i.e. the large standard deviations reported in Table 4). There are two kinds of variability that are important for our purposes: Variability in the relative *weighting* of each of the forgiveness cues and variability in the relative *direction* of that weighting. That is, two people can weight some forgiveness cue with the same magnitude weighting, but one person might weight it *positively* such that it is associated with a judgement that forgiveness has occurred, while the other person might weight it *negatively* such that it is associated with a judgement that forgiveness *has not* occurred. Thus, there are two separate issues to examine: (1) the extent to which each forgiveness cue is weighted in an *absolute* sense, and (2) the direction of the relationship between forgiveness judgments and the forgiveness cue weightings across participants.

First, we examined whether some forgiveness cues are, on average, weighted more heavily than others. To do this we first took the absolute value of the standardised beta weights for each participant and analysed them using a repeated measures ANOVA with a between-subjects factor of valence (low, high) and a within-subjects factor of forgiveness cue (~nE, ~rA, cG, ~pT, bR, and R). The results of the ANOVA found a significant main effect of forgiveness cue F(5, 960) = 38.613, *p* < .001 (see Figure 1). The significant main effect of forgiveness cue indicates that participants, overall, attached more importance to some cues than to others in making their forgiveness judgments. Further, in line with a previous observation of the descriptive data, there was no significant main effect of valence, and no significant interaction between valence and forgiveness cue.

*Figure 1. Overall absolute standardised beta weights for each forgiveness cue. Error bars represent standard error.*

Pairwise comparisons using a Bonferroni correction were carried out on the main effect of forgiveness cue. First, the absolute standardised beta coefficient of ~rA (M = .33, *S.D.* = .25) was significantly larger than ~nE (M = .24, *S.D.* = .19; *p* < .001), cG (M = .10, *S.D.* = .08; *p* < .001), ~pT (M = .23, *S.D.* = .21; *p* = .002), and R (M = .10, *S.D.* = .14; *p* < .001). There was no significant difference between the absolute standardised beta coefficient of ~rA and bR (M = .31, *S.E.* = .29; *p* > .999).

Second, the absolute standardised beta coefficient of bR was significantly larger than cG and R (*p* < .001). There was no significant difference between the absolute standardised beta coefficient of bR and ~nE (*p* = .325), ~rA (*p* > .999), or ~pT (*p* = .085).

Third, the absolute standardised beta coefficient of ~nE was significantly larger than both cG (*p* < .001) and R (*p* < .001). There was no significant difference between the absolute standardised beta coefficient of ~nE and ~pT (*p* > .999), or bR (*p* = .325).

Fourth, the absolute standardised beta coefficient of ~pT was significantly larger than both cG and R (*p* < .001). There was no significant difference between the absolute standardised beta coefficient of ~pT and ~nE (*p* > .999), or bR (*p* = .085).

Finally, there was no significant difference between the absolute standardised beta coefficient of cG and R (*p* > .999).

Second, we examined the relative direction of the weighting of each forgiveness cue. To do this we determined the proportion of people whose standardised beta coefficient for each forgiveness cue were positive compared to negative (or 0).[[25]](#footnote-25) These proportions, as well as corresponding tests for differences between high and low valence conditions, are presented in Table 5. Table 5 also shows that the large majority of people weight the forgiveness cues positively, except for ~pT, which only a small minority of people weight positively. Further, from these tests we can see that there is no association at all between valence and the overall relative direction of the weighting of the forgiveness cues.

*Table 5. Proportions of participants displaying positive standardised beta weights to each forgiveness cue. Comparison is between low valence and high valence condition.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Forgiveness Cue** | **Low Valence (n = 93)****%** | **High Valence (n = 101)****%** | ***X2*** | ***p*** |
| ~nE | 77.4 | 73.3 | .448 | .503 |
| ~rA | 89.2 | 84 | 1.493 | .222 |
| cG | 71 | 74.3 | .264 | .607 |
| ~pT | 10.8 | 13.9 | .432 | .511 |
| bR | 86 | 85.1 | .030 | .863 |
| R | 67.7 | 67.3 | .004 | .951 |

*Note: DF = 1 and N = 194, for all X2 tests.*

**4. In Favour of Folk Conceptual Pluralism**

First, we find relatively high weightings given to cues ~nE, ~rA, ~pT, and bR. This shows that participants, by and large, think that letting go of anger (~nE) and resentment, (~rA), and believing that the perpetrator has repented (bR) are more important than other forgiveness cues when determining whether a scenario is one in which forgiveness occurs. Since many philosophers have thought these three cues to be important, this goes at least some way towards vindicating some of their intuitions. However, contrary to what philosophers have thought, the victim not intending to punish the perpetrator (~pT) counts *against* a scenario being one in which forgiveness has taken place.

Yet these results clearly do not vindicate any of the core accounts of forgiveness theoreticians have offered. For instance, while Griswold (2007) takes a belief in repentance to be necessary for forgiveness, he also thinks that actual repentance is required. Yet we found that participants placed relatively low weight on cue R (repentance), contra Griswold. Similarly, Boleyn-Fitzgerald supposes that removing anger is sufficient for forgiveness, While participants did heavily weight the removal of anger, (~nE) they also heavily weighted the removal of resentment (~rA), and the belief that there is repentance (bR). Hence, viewed as a descriptive claim about the folk concept, Boleyn-Fitzgerald’s view is not borne out by our data.

While many philosophers take the letting go of resentment to be necessary for forgiveness, and our results bear out that participants did highly weight this cue, most of these theorists have supposed that other factors are also necessary. For instance, Garrard and McNaughton (2002), and Hampton (1988) all take cognition (cG) to also be necessary; yet our results found that participants do not weight this cue highly. Similarly, while Couto (2016) and Hieronymi (2001) both take the removal of resentment to be necessary for forgiveness, they also suppose that cognition (cG) and repentance (R) are necessary: but, once again, we found both of these cues received very low weightings. Finally, cue 4 (~pT) was *negatively* weighted. That is, participants were more likely to say that a scenario was one in which forgiveness occurred, if it was one in which the victim *still intended to* punish the perpetrator, rather than forgoing punishment. This latter result contradicts Zaibert’s (2009A) suggestion that declining to punish weights in favour of, as opposed to against, there being forgiveness.

Having said this, although our results show that some cues are more highly weighted than others, our results suggest that there is no single, univocal, folk concept that is being deployed by participants. First, while we see no difference between results in the high valence and low valence conditions, we see very large variance between participants. In particular, when we look at the standard deviations for each cue, we see that in most cases this is *larger than the mean itself*. This reflects massive differences in participants’ responses. It means that for each cue that is positively weighted, for some participants that cue is negatively weighted (and *vice versa* for ~pT). Another way to express this is that for each cue, there is a very significant percentage of the population for whom that cue is important, and another significant percentage for whom it is not. As noted above, the cue of giving up resentment was found to be important to the majority of participants (62.9%), nevertheless leaving nearly 40% of participants for whom it was not important. Likewise, giving up anger (52.6%), believing the perpetrator has repented (49%), and intending not to punish (50.5% were important to only half of the participants, and hence not important to the other half. Finally, the victim working through their negative emotion and reactive attitudes and identifying their source (17.5%) and the perpetrator being remorseful (12.9%) were important to only a minority of participants.

One might think that the divergent responses are due to the absence of a verbal declaration of forgiveness. However, when asked directly participants overwhelmingly (86.1%) report that Jane does not have to let Phoebe know that she is forgiven in order for forgiveness to have occurred.[[26]](#footnote-26)

One of the weaknesses of our study is that we only consider how participants respond to the presence and absence of each individual forgiveness cue. Maybe what’s important for people’s forgiveness judgements is not just the presence or absence of individual forgiveness cues but the various interactions between those cues. Unfortunately, this study lacked the statistical power to test all these interactions, something future studies might address. Nevertheless, given the importance that philosophers place on each of these cues singly, as well as in concert, we would expect each of them to be associated with forgiveness judgements independent of one another.

These divergent responses constitute *prima facie* evidence in favour of folk conceptual pluralism about forgiveness. If we thought that we could evaluate philosophical accounts of forgiveness by asking whether they accord with the unified folk concept, we must think again, for it appears that there is no unified folk concept. Indeed, the evidence here suggests that even those theorists who countenance the existence of folk conceptual pluralism may not have gone far enough: the very divergent responses that we see suggests that there is more than just one or two concepts deployed in the population in question.

In fact, this evidence is supported by recent work on the concept of forgiveness undertaken by psychologists. These studies purport to show evidence in favour of a prototype view (Friesen and Fletcher 2007; Kearns and Fincham 2004) by showing converging responses amongst subjects asked to consider what it means to forgive. In fact though, *relatively* little convergence is discovered. Friesen and Fletcher (2007) report that 51.8% of folk surveyed mentioned that forgiveness involves being “relieved, like a weight or burden has lifted”, 37.3% mentioned a return to “general everyday communication”, 33.6% mentioned a “hug or embrace between victim and perpetrator”, 33.2% mentioned consideration of the “personal consequences of forgiving or not forgiving” and 32.3% mentioned “feeling happy or glad”. A similar study, conducted by Kearns and Fincham (2004), asked participants how they would describe forgiveness to someone unfamiliar with the concept. They found a modest correlation between the features listed by participants in each study (r = .61, *p* < .001). Far from showing that forgiveness is a prototype concept, of which some of these features are central to that prototype, we think that, in combination with our results, this suggests a fair degree of divergence regarding what constitutes forgiveness.

How might forgiveness theorists respond to this evidence? They might resist folk conceptual pluralism by arguing that in fact forgiveness is a functional concept, or they might accept folk conceptual pluralism, and turn their attention to the question of which of the concepts in this domain (or others yet to be engineered) we have reason to deploy. We consider each of these options in turn.

**4.1 Forgiveness as a Functional Concept**

Here is one way the forgiveness theorist might resist endorsing folk conceptual pluralism. She might argue that there is a single, univocal folk concept of forgiveness, which is a higher-level *functional* concept. As such, the content of the concept will be something like: forgiveness is whatever combination of mental states and behaviours play functional roles F1…Fn. Perhaps F1…Fn specifies a whole set of psychological and social roles that something needs to play if it is to count as an instance of forgiveness. Perhaps it is these roles to which the folk have (tacit) access. Then theorists and folk alike might *agree* that forgiveness is whatever combination of mental states and behaviours play functional roles F1…Fn but disagree about which mental states and behaviour in fact realise those roles.

For instance, it could be suggested that forgiveness plays the functional role of removing conflict between victim and perpetrator.[[27]](#footnote-27) If this were true, and if some of the folk believe that the perpetrator’s remorse removes the conflict, whereas other folk believe that the victim’s loss of resentment removes the conflict, then the apparent disunity revealed in our results might sit beneath a higher-level conceptual unity. Alternatively, it might be suggested that forgiveness has been achieved whenever the victim no longer needs to do anything about the wrong; the victim has forgiven the perpetrator if and only if the victim is free to move on. If this were true, and if some of the folk believe that the victim is free to move on whenever she has cognitively processed the wrong, whereas other folk believe that the victim is free to move on only after the perpetrator has been punished and has shown remorse, then the apparent disunity in our experimental results is compatible with there being a unified folk concept of forgiveness. Quite generally, if our participants agree about the role(s) that forgiveness plays, but disagree about which scenarios describe circumstances in which something plays that role, then the divergence is not best explained by their deploying different concepts. Instead, it is evidence of disagreement about the conditions under which the relevant set of roles are realised.

In order to test whether any of these roles underpins the pattern of folk judgements observed in the present study, we would need to devise further empirical research to probe subjects’ views about the functional roles of forgiveness. However, we are sceptical that such research will find convergence about which roles are the relevant ones. If there is widespread agreement regarding what the roles are, then it is perplexing that there is as much disagreement as there is about which things fall under the concept. In other words, if there is agreement as to the roles forgiveness plays, then we should expect greater convergence regarding when forgiveness obtains, and when it does not, even if there are some difficult cases in which there could reasonably be dispute about whether the relevant roles are being played. Nevertheless, we think that investigation of this option is important, and that our data does not rule it out.

**4.2 The Conceptual Ethics of Forgiveness**

A second option is for forgiveness theorists to accept that folk conceptual pluralism is true, and abandon any attempt to articulate the content of a shared folk concept. They might instead aim to catalogue the range of varied folk conceptions, engineer new concepts in this domain, evaluate all of the contenders, and then promulgate the best of these concepts. This revisionary project would be an example of what is known as conceptual ethics.[[28]](#footnote-28)

One variant of conceptual ethics is methodologically modest. It seeks to determine whether there are non-moral reasons to prefer some of the available concepts of forgiveness while rejecting others. For instance, it could turn out that some are internally inconsistent, and hence inferior to their rivals. Alternatively, we might discover that some conceptions of forgiveness are a comparatively bad fit with our best theories of related phenomena, such as resentment, punishment, and mercy, and hence ought to be rejected. Perhaps we will find that some of the available concepts of forgiveness do a better job of unifying a disparate set of intuitions, and possess a kind of simplicity and explanatory power that other conceptions lack. In favouring one concept over the others on these grounds, we would be moving beyond folk conceptual pluralism to defend a unified account of forgiveness, while admitting that our favoured account will not map onto the outlines of the (non-existent) unified folk concept of forgiveness.

A second variant of this project is more methodologically radical. We could try to identify or design the concept of forgiveness that best allows us to achieve certain moral or social goals, and reject competing conceptions of forgiveness that are not as useful in this respect (Boleyn-Fitzgerald 2002). Let us call this morally reformative conceptual ethics.

Theorists who attempt morally reformative conceptual ethics will likely have different moral goals. Here we identify four broad goals conceptual ethicists might emphasise. This is not intended to be an exhaustive list, but rather simply to emphasise that different goals will likely pull in different directions. Firstly, conceptual ethicists might seek whichever concept best helps victims find inner peace. We could expect such a concept to focus on the psychology of the victim, since such a concept would allow victims to forgive and move on even when the perpetrator is staunchly unrepentant. We could call concepts of this kind *victim focused.* Secondly, if our moral goal were to relieve perpetrators of disproportionate guilt and suffering we would expect to recommend concepts that are focused on the perpetrator’s psychologies. We could call these concepts *perpetrator focused.* Thirdly, if one’s moral goal is to bring about certain inter-personal changes between victim and perpetrator, then one would likely recommend a *relational concept* that focuses on the psychological features of both, and that characterises forgiveness as communicative. Finally, if our moral goal is to achieve broader social benefits, we might recommend a *community focused* concept of forgiveness that has whatever features are most likely to promote social healing and to minimise the risk of repeat offences.

Thus it is clear that morally reformative conceptual ethicists must answer some difficult questions. Which moral goal or goals ought we pursue, and when these goals conflict with each other, how should the conflict be resolved? Helping victims find inner peace is morally laudable, but so too is the goal of helping perpetrators remove disproportionate guilt, *and* the goal of bringing unrepentant perpetrators to justice, *and* the goal of minimising the risk of future wrongdoing, *and* the goal of healing social rifts, *and* the goal of fostering self-respect. Moreover, these goals pull us in different directions: we have no reason to suppose that a concept of forgiveness best facilitates social healing will also best promote justice or self-respect.

We predict that disagreements over which specific moral goals conceptual ethicists ought to adopt will be very tough to resolve. For example, whilevictim focused concepts seem best suited to therapeutic settings, the fact that they make no requirements on the perpetrator means that they are not well placed to do the work desired by those who prioritise social goals, including deterrence. Similarly, while relational conceptions are best suited to the mending of broken relationships, such conceptions are not ideal for bringing inner peace to victims, and might undermine attempts to secure social healing or justice. On the other hand,community focused conceptions are best suited to the context of building social policy and ensuring justice, and may do little to assist in the psychological healing of either victims or perpetrators, rendering such concepts unhelpful in a therapeutic setting.

Morally reformative conceptual ethicists also ought to ask what evidence there is that philosophical advocacy of an engineered concept of forgiveness will actually lead to the deployment of that concept outside the academy, and to the broad moral or social benefits that are our supposed goal. We ought to be wary of wishful thinking in this domain. There are venerable and influential traditions and practices—religious, legal, and cultural—which influence the way that the folk think about forgiveness. These practices may be very hard to shift.

Our empirical research into the folk conceptions can provide some guidance in evaluating the feasibility of projects in conceptual ethics. Our research does not, of course, tell us how many concepts are present in the population we sampled, nor what the content is, of each of those concepts. In order fully to consider issues of feasibility we would need to know what content these concepts have, and in particular, to know which conditions must, of necessity, obtain for any particular concept to be satisfied. Nevertheless, we can provide some guidance on this issue in the following sense: in the absence of firm data regarding the content of these concepts, it seems reasonable to suppose that where we find a substantial percentage of the population weighting a particular cue heavily, this gives us some reason (albeit defeasible) to suppose that an engineered concept that does not include this cue may be too distant from most of the concepts currently deployed, to be able to catch on amongst the folk. This is, of course, defeasible reason, because it does not follow that just because some cue is heavily weighted by a large percentage of the population, that some, or all, of that population takes the presence of that cue to be a necessary condition for their concept to be satisfied.

Still, bearing in mind the limitations here, we can at least reflect on some defeasible reasons our study provides to be wary of certain engineering projects which recommend concepts which do not include, as part of their content, cues that are heavily weighted by a substantial percentage of the population. For in such cases we do have defeasible reason to worry that such engineered concepts might be very difficult to inculcate in the population. For example, participants highly weighted the giving up of resentment (~rA). Thus, although this condition may not be required in community-focused conceptions of forgiveness, it may be infeasible for conceptual engineers to advocate a concept that eschewed this feature.

Similarly, participants highly weighted the belief that the perpetrator has repented (bR). This suggests that Boleyn-Fitzgerald’s (2002) conceptual engineering project may be too ambitious: the folk may not be able to come to deploy a concept of forgiveness which requires only that the victim releases her negative emotions. More generally, this suggests that victim-focused concepts might be infeasible to deploy, and therefore the goal of purely therapeutic forgiveness may be one that cannot be pursued in isolation.

Finally, as with all social engineering projects, well-meaning interventions in folk discourse about forgiveness may have unforeseen and undesirable side-effects. While the more radical morally reformative conceptual ethics project is both promising and exciting, forgiveness theorists who pursue it must engage directly with these difficult issues.

**5. Conclusion**

Descriptive theorising about forgiveness has reached an impasse, with theorists standing firmly by their divergent intuitions. Our data suggests that these disputes cannot be resolved by appeal to single unified folk concept of forgiveness, because there is no such concept. Thus, we hope to have shown the way forward in debates about forgiveness. Descriptively, there is work to be done in spelling out the apparent plurality of folk concepts, and in identifying various functional roles that forgiveness might play. Normatively, there is work to be done identifying a concept which is not too distant from the folk concept(s), but which we have good reasons—moral or otherwise—to prefer.

**References**

Allais, Lucy, 2013, 'Elective Forgiveness', *International Journal of Philosophical Studies*, 21 (5), 637-53.

–––, 2008, ‘Wiping the Slate Clean: The Heart of Forgiveness,’ *Philosophy and Public Affairs*, 36(1): 33–68

Boleyn-Fitzgerald, Patrick, 2002, “What Should ‘Forgiveness’ Mean?” *Journal of Value Inquiry*, 36: 483–498.

Boon, S.D, and Sulsky, L. M., (1997). “Attributions of Blame and Forgiveness in Romantic Relationships: A Policy-Capturing Study”. *Journal of Social Behaviour and Personality.* 12 (1): 19-44.

Burgess, Alexis & Plunkett, David (2013A). ‘Conceptual Ethics I.’ *Philosophy Compass* 8(12):1091-1101.

Burgess, Alexis & Plunkett, David (2013B). ‘Conceptual Ethics II.’ *Philosophy Compass* 8(12):1102-1110.

Burgess, Alexis, Cappelen, Herman & Plunkett, David. (forthcoming). *Conceptual Engineering and Conceptual Ethics*. OUP.

Butler, Joseph, 1726/2002, ‘Upon Resentment and Forgiveness of Injuries’, *Sermons VIII*, (Cambridge: Hilliard and Brown).

Chalmers, D. and Jackson, F. (2001). ‘Conceptual Analysis and Reductive Explanation’, *Philosophical Review,* 110 (3):315-61.

[Chalmers, D. J. (2012). *Constructing the World.*](https://philpapers.org/rec/JCHCTW) Oxford: Oxford University Press.

Couto, Alexandra (2016). ‘Reactive Attitudes, Forgiveness, and the Second-Person Standpoint’. *Ethical Theory and Moral Practice* 19: 1309-1323.

Floridi, Luciano (2013). *The Ethics of Information*. Oxford: OUP.

Friesen, M.D. & Fletcher, G. J. O. (2007). Exploring the lay representation of forgiveness: Convergent and discriminant validity. *Personal Relationships*, *14*, 209–223.

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

Govier, Trudy, and Wilhelm Verwoerd, 2002, “Forgiveness: The Victim's Prerogative,” *South African Journal of Philosophy*, 21: 97–111.

Green, S. B. (1991). How Many Subjects Does It Take To Do A Regressions Analysis? *Multivariate Behavioral Research*. 26(3): 499-510.

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

–––, 2009, ‘Forgiveness, Secular and Religious: A Reply to My Critics’, *Proceedings of the American Catholic Philosophical Association*, 82: 303-313.

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

Hampton, Jean, 1988, ‘Forgiveness, Resentment, and Hatred,’ in *Forgiveness and Mercy*, Hampton, Jean, and Jeffrey G. Murphy (eds.), Cambridge: Cambridge University Press, pp. 35–87.

Hieronymi, Pamela (2001). ‘Articulating an uncompromising forgiveness.’ *Philosophy and Phenomenological Research*, 62(3):529-555.

Holmgren, Margaret, 1993, ‘Forgiveness and the Intrinsic Value of Persons’, *American Philosophical Quarterly*, 30(4), 341-352.

–––, 2012. *Forgiveness and Retribution: Responding to Wrongdoing.* Cambridge University Press.

Jackson, Frank, 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, London: Oxford University Press.

Kearns, J. N., & Fincham, F. D. (2004). A prototype analysis of forgiveness. *Personality and Social Psychology Bulletin*, *30*, 838–855.

Knobe, J. and S Nichols (2008). *Experimental Philosophy*. OUP.

Kolnai, Aurel, 1973-1974, 'Forgiveness', *Proceedings of the Aristotelian Society* 74: 91-106.

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

–––, 1988, ‘Forgiveness, Mercy, and the Retributive Emotions’, *Criminal Justice Ethics*, 7 (2): 3–14.

–––, 2001, ‘Forgiveness’, in Lawrence C. Becker and Charlotte B. Becker (eds.), *Encyclopedia of Ethics*. New York: Routledge: 561–562.

Murphy, Jeffrie G. & Jean Hampton, 1988, *Forgiveness and Mercy*, Cambridge: Cambridge University Press.

Novitz, David, 1998, ‘Forgiveness and Self-Respect’, *Philosophy and Phenomenological Research*, 58 (2): 299–315.

Pettigrove, Glen, 2004, ‘The Forgiveness We Speak: The Illocutionary Force of Forgiving’, *The Southern Journal of Philosophy*, 42(3): 371–392.

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

Plunkett, David, 2015, ‘Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy’, *Inquiry* 58 (7-8):828-874

–––, 2016, ‘Conceptual History, Conceptual Ethics, and the Aims of Inquiry: A Framework for Thinking About the Relevance of the History/Genealogy of Concepts to Normative Inquiry’, *Ergo, an Open Access Journal of Philosophy* 3.

Prusak, Bernard, 2008. ‘What Are the ‘Right Reasons’ to Forgive? Critical Reflections on Charles Griswold’s Forgiveness: A Philosophical Exploration,’ *Proceedings of the American Catholic Philosophical Association*, 82:287-295

Russell, Luke, 2016, ‘Forgiving While Punishing’, *Australasian Journal of Philosophy*, 94(4):704-718.

Richards, Norvin, 1988, ‘Forgiveness’, *Ethics* 99:77–97.

Strawson, Peter F. (1962). Freedom and resentment. In Gary Watson (ed.), *Proceedings of the British Academy*, Volume 48. OUP: Oxford: 1-25.

Warmke, Brandon, 2013, ‘Two Arguments Against the Punishment-Forbearance Account of Forgiveness’, *Philosophical Studie*s, 165(3): 915–920.

–––, 2011, ‘Is Forgiveness the Deliberate Refusal to Punish?’, *Journal of Moral Philosophy*, 8(4): 613–620.

–––, 2014, ‘Forgiveness and Responsibility’ Dissertation, <http://hdl.handle.net/10150/333086>

–––, 2016, ‘The Normative Significance of Forgiveness’, *Australasian Journal of Philosophy*, 94(4): 687-703

Zaibert, Leo, 2012, ‘On Forgiveness and the Deliberate Refusal to Punish: Reiterating the Differences’*, Journal of Moral Philosophy*, 9: 103–113.

–––, 2009A, ‘The Paradox of Forgiveness’, *Journal of Moral Philosophy*, 6: 365–393.

–––, 2009B, ‘Forgiveness: An Introduction’, *The Monist*, 92(4): 481–488.

1. See Jackson (1998) for a discussion of conceptual analysis and folk concepts. [↑](#footnote-ref-1)
2. One notable exception is Warmke (2014), who does appeal to intuitions evoked by simple thought experiments, but who argues for a prototype view of forgiveness rather than trying to identify necessary or sufficient conditions. [↑](#footnote-ref-2)
3. e.g. Griswold (2007) and Zaibert (2009A). [↑](#footnote-ref-3)
4. e.g. Pettigrove (2012), Hieronymi (2001), Garrard and McNaughton (2002), Allais (2008) and Russell (2016). [↑](#footnote-ref-4)
5. When Griswold is taken to task by Prusak (2008:288), regarding whether his analysis of forgiveness “does justice to the ways that people speak of and practice forgiveness”, he simply points out that no analysis will do so, as “people speak of and practice forgiveness in ways inconsistent with each other” (2009:308). However, he still contends that his theory “does honor at least some of our most important convictions, and is maximally inclusive.” (2009:308). [↑](#footnote-ref-5)
6. Warmke (2016) claims that this speech act is at the heart of the central, paradigmatic cases of forgiveness, but is open to the idea that peripheral cases of forgiveness do not include all the central features. [↑](#footnote-ref-6)
7. In this, we follow Boon and Sulsky (1997) who ran a policy-capturing study to examine how people weight and combine various factors about the nature of a transgression to draw conclusions about a romantic partner’s culpability and their own willingness to forgive. [↑](#footnote-ref-7)
8. See for instance Chalmers and Jackson (2001) and Chalmers (2012). [↑](#footnote-ref-8)
9. **~**nE should not be read simply as it not being the case that the victim experiences negative emotions as the result of the perpetrator’s wrongdoing but, rather, as the cue according to which, if the victim experiences negative emotions as the result of the perpetrator’s wrongdoing, then the victim has let go of those negative emotions. [↑](#footnote-ref-9)
10. As with ~nE, ~rA should be read as the cue according to which, if the victim experiences negative reactive attitudes as the result of the perpetrator’s wrongdoing, then the victim has let go of those reactive attitudes. [↑](#footnote-ref-10)
11. Boleyn-Fitzgerald (2002) argues that we ought adopt a conception of forgiveness according to which ~nE is both necessary and sufficient. [↑](#footnote-ref-11)
12. Many philosophers take ~rA to be necessary for forgiveness, including *inter alia*, Allais (2008) and Holmgren (1993). [↑](#footnote-ref-12)
13. Whilst no theorists take cG to be sufficient for forgiveness, it is common to suppose that it is necessary. See i*nter alia* Novitz (1998) and Griswold (2007). [↑](#footnote-ref-13)
14. Zaibert (2009A) takes refusing to punish the perpetrator (for certain reasons) to be necessary for forgiveness, and, when combined with the persistence of the culpability judgments, sufficient. [↑](#footnote-ref-14)
15. Griswold (2007) takes bR to be necessary for forgiveness. [↑](#footnote-ref-15)
16. Griswold (2007) takes R to be necessary for forgiveness. [↑](#footnote-ref-16)
17. Those who take ~nE or ~rA to be sufficient will take this combination of cues to be sufficient. Those who take *both* ~nE and ~rA to be necessary will take this combination of cues to be necessary. Richards (1988) takes this combination to be both necessary and, in conjunction with the persistence of the culpability judgments, sufficient. [↑](#footnote-ref-17)
18. Garrard and McNaughton (2002) and Hampton (1988) take this combination, in conjunction with the persistence of the culpability judgments, to be necessary sufficient for forgiveness. [↑](#footnote-ref-18)
19. Griswold (2007) takes this combination condition to be necessary for forgiveness. [↑](#footnote-ref-19)
20. Couto (2016) takes this combination to be necessary for virtuous forgiveness. [↑](#footnote-ref-20)
21. This is the combination of all the cues we tested. Hieronymi (2001) comes very close to saying that this combination is necessary and, in conjunction with the persistence of the culpability judgements, sufficient for forgiveness, but in a footnote she says she is open to the view that broad communal condemnation of the perpetrator’s wrong action could substitute for R and bR. [↑](#footnote-ref-21)
22. The capacity to take the test has been temporarily disabled, since the only way to do so requires seeing who has ethics clearance for the project, and hence means the paper is no longer blinded. For our reviewers’ benefit, the complete set of vignettes is included in a separate attached document. [↑](#footnote-ref-22)
23. This count simply reflects whether or not a participant possessed a religious belief or not. Given the wide array of religious beliefs present in the sample and our current sample size we are unable to accurately assess or account for possible interfaith differences. [↑](#footnote-ref-23)
24. R-squared is a goodness of fit statistic. A value of 0 indicates that the model explains nothing, while a value of 1 indicates the model explains all the variability of the response data. [↑](#footnote-ref-24)
25. The number of beta weightings of 0 for each forgiveness cue are as follows: ~nE (7), ~rA (6), cG (6), ~pT (3), bR (6), and R (9). [↑](#footnote-ref-25)
26. X2(1) = 101.031, *P* < .001. It seems that the folk do not agree with Warmke’s (2016) and Griswold’s (2007) communicative accounts of forgiveness. [↑](#footnote-ref-26)
27. Presumably this would be just one role forgiveness would need to play, since excusing, or forgetting, could also play this role. [↑](#footnote-ref-27)
28. For discussion, see Burgess, Cappelen and Plunkett (forthcoming), Floridi (2011), and Burgess and Plunkett (2013A, 2013B). [↑](#footnote-ref-28)