

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

"I forgive you!" — Exploring the Impact of Forgiveness on Negative Emotions and Blame

Permalink

<https://escholarship.org/uc/item/4pz2v8m7>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Author

Willemsen, Pascale

Publication Date

2023

Peer reviewed

“I forgive you!” —

Exploring the Impact of Forgiveness on Negative Emotions and Blame

Walker Siu Hong Au (SiuHong.Au@uzh.ch)

University of Zurich, Department of Philosophy, Zollikerstrasse 117, 8008 Zurich, Switzerland

Pascale Willemsen (Pascale.Willemsen@uzh.ch)

University of Zurich, Department of Philosophy, Zollikerstrasse 117, 8008 Zurich, Switzerland

Abstract

Forgiveness plays a significant role in our everyday social life, and, because of that, it has received an increasing amount of attention in academic research. However, philosophers and psychologists are equally worried by the fact that we still lack an empirically adequate characterization of forgiveness. In this paper, we present two preregistered studies in which we explore what ordinary people believe a speaker does when he or she performs the speech act of forgiving by uttering the phrase, “I forgive you”. Study 1 uses a vignette-based stimulus to examine what participants believe to change after the victim granted forgiveness to their wrong-doer. In Study 2, we apply a linguistic test, the cancellability test, to determine whether participants consider forgiving the wrong-doer but still blaming them compatible.

Keywords: blame; cancellability test; emotion accounts to forgiveness; forgiveness; restorative justice; vignette study

Introduction

Forgiveness has strong religious connotations. It is considered one of the major virtues or even duties in Islam, Christianity, Judaism, Buddhism, and other religious and spiritual movements. For instance, in Mishneh Torah, Teshuvah 2:10 reads: “It is forbidden to be obdurate and not allow yourself to be appeased. [...] When asked by an offender for forgiveness, one should forgive with a sincere mind and a willing spirit.” In the Bible, Luke 6:37 says, “Do not judge, and you will not be judged. Do not condemn, and you will not be condemned. Forgive, and you will be forgiven.”

But even in other aspects of life, the meaning of forgiveness cannot be overestimated. Iranian law gives the family of murder victims the option to grant forgiveness to the offender, thereby preventing them from being executed¹. In contrast, the American criminal juridical system does not consider whether the victim grants forgiveness as a mitigating factor in sentencing.² Yet, many scholars have argued that a juridical system assigning greater relevance to *restorative justice* would be preferable (Bibas, 2006; Gehm, 1992; Lacey & Pickard, 2015; Lerman, 1999; Nygaard, 1997). Experimental evidence confirms that retributive and restorative justice are not just compatible but that a balanced

combination benefits victims, their families, the offenders, and society as a whole (Shapland, 2016; Strelan et al., 2008; Witvliet et al., 2008). The conciliatory effect of forgiveness on society is further documented by the success of restorative justice programs, for instance, in reaction to the abolition of apartheid in South Africa or the Rwandan genocide.

While forgiveness has only played a subordinate role in early modern psychology (notable exceptions being (Emerson, 1964; Heider, 1958; Litwinski, 1945; Piaget, 1932; Rokeach, 1973), it has recently enjoyed a significant increase in empirical attention (Boon & Sulsky, 1997; Darby & Schlenker, 1982; McCullough et al., 2003; Weiner et al., 1991), especially with the rise of positive psychology in 1998. Surprisingly though, psychologists remain divided about the most appropriate definition of forgiveness. Psychologist Everett Worthington (2005) argued that determining the proper definition of forgiveness was the “major issue characterizing this new science of forgiveness”, while McCullough (2000, 7) criticizes that “no consensual definition of forgiveness exists”. This lack of conceptual clarity is not unique to psychological research but also a central issue in the philosophical debate (see Hughes & Warmke, 2022) for a discussion.

Philosophers and psychologists alike (e.g., Bennett, 2018; Helmreich, 2022; Warmke, 2016) have recently emphasized the interpersonal, social dimension of forgiveness and argued that it is crucially about repairing social relationships between the victim and the wrong-doer. It is often expressed verbally, overtly, and publicly—at least between the victim and wrong-doer. Mutual expectations change with the social act of verbally expressing one’s forgiveness. For example, we can expect a victim to have (at least partly) overcome their negative feelings for the wrong-doer. In addition, after the victim forgives, she no longer expects apologies and restitution from the wrong-doer (Warmke, 2016).

It is interesting that, upon hearing a victim express their forgiveness, we can make inferences about how they feel and what they expect. It might be argued that inferences like these do not just inform us about the nature of the *social practice* of forgiving. It further allows for deeper insights into *forgiveness itself*. Joram Haber and others have recently argued that the question “What is forgiveness?” needs to be

¹ The legal and ethical aspects of this practice are, of course, heavily debated, especially since Iranian laws draw the justification of this practice from Islam. The point we wish to make is merely that forgiveness plays a role in some judicial systems.

² See McThenia (1999) for a discussion of the different ways in which forgiveness can still affect sentencing in capital murder cases under Anglo-American law.

answered “in the context of what speakers mean when they employ the term” (Haber, 1991, p. 53; see also, e.g., Austin, 1975; Nelkin, 2013; Pettigrove, 2004). More specifically, Haber suggests investigating what a person does when they perform the speech act of forgiving by uttering phrases such as, “I forgive you”.

To this end, we conducted two preregistered experiments on what people communicate when they grant forgiveness. In our first study, we presented participants with one of four stories in which a wrong-doer violated a social or moral norm and later asked for forgiveness. We compared the effect of the victim granting forgiveness on a series of cognitive, conative, and affective information that participants inferred from the situation. Our results indicate that people believe that with forgiveness being granted, the victim’s attitude, especially their emotions and behavioral dispositions, have changed. In a second experiment, we followed this up by presenting participants with cancellability tests. This second experiment supports the idea that while the victim feels significantly less negatively about the wrong-doer, negative feelings can remain even after forgiveness is granted. We discuss how our results shed light on the fundamental question of what forgiveness is.

What Changes When We Forgive?

The dominant view in philosophy and psychology, the Emotion Account of forgiveness, states that forgiveness is essentially a matter of how one *feels* about one another (e.g., Hughes, 1993; Moore, 1989; Murphy, 1988). This suggestion seems to match ordinary intuitions. We associate forgiveness with overcoming, eliminating, or forswearing various negative feelings, such as anger, frustration, or resentment. When we ask for forgiveness as a wrong-doer ourselves, we often mean to ask the victim to no longer be angry with us.

However, philosophers and psychologists have remained rather vague and non-committal concerning the question of *what specific negative emotions change when we forgive*. While exemplary lists of emotions are often provided, it is not clear whether changes in these emotions are necessary for someone to forgive. Philosophers have proposed a list that contains, among other items, *resentment, anger, hatred, loathing, contempt, indifference, disappointment, and sadness* (Darwall, 2006; Hughes, 1993; Hughes & Warmke, 2022; Murphy, 2003). Most philosophical attention has been paid to the feeling of resentment—only to leave the question of what resentment is unanswered. In psychotherapy, the negative emotions under consideration include, e.g., *resentment, bitterness, hostility, hatred, anger, fear, and bitterness* (Enright & Fitzgibbons, 2015; Worthington Jr. & Sandage, 2016).³

³ Going beyond psychological research and psychotherapy, in the legal debate, it has been suggested that “forgiveness is the private, emotional response of a victim” (Allen, 2015) in which the victim “cancels a debt” or decides to “no longer hold a grudge” and shows love and compassion for the wrong-doer (Strelan et al., 2008).

While all these suggestions seem plausible candidates for a theoretical, academic discussion, “hatred” or “loathing” might have too extreme connotations in comparably mundane cases. Suppose someone blabbed about a minor secret you told them. While this behavior violated a norm and an apology is in order, many people might consider it overly dramatic to say that before forgiving, you hate or loathe the wrong-doer.⁴

For our studies, we chose to use a selection of those emotions discussed in the literature that seemed most appropriate in the situational contexts we provide: anger, sadness, and disappointment. In addition to testing whether granting forgiveness affects those emotions in any way, *we also test whether participants believe the negative emotions are entirely overcome*.

While forgiveness is thought to involve a change in one’s emotional or affective attitudes, it does not seem to necessarily involve a change in beliefs about the moral status of the action. It seems natural to believe that the action for which forgiveness was granted remains wrong or bad, and that the victim still believes the behavior crossed a line and justified a negative response. This seems to be particularly indisputable in cases of legal relevance. It would be absurd to suggest that the family of a murder victim has changed their mind about the wrongness of the murder just because they forgave. Most philosophers seem to find such intuitions quite compelling and agree that while our affective attitudes towards the wrong-doer *necessarily* change, our beliefs about the wrongness of the act *can, but do not have to change* (Allais, 2008, 2013; Brunning & Milam, 2022; Hieronymi, 2001).

While, *prima facie*, it seems plausible that some negatively valenced beliefs about the action’s moral status are compatible with forgiveness, we currently lack empirical evidence of this in ordinary people. Therefore, we ask *what beliefs exactly are consistent with a victim’s sincere forgiveness*. The most promising candidates for our investigation seem to be the beliefs that the action was *wrong*, the agent is *blameworthy* for it, and that they are *at fault* for the harm they caused.

Following suggestions by Brunning and Milam (2022), we also consider reactions closely connected to behavioral dispositions and expectations. For one thing, we consider whether *blame* is a central element in forgiveness. Both philosophers (e.g. Allais, 2008, 2013; Brunning and Milam, 2022; Hughes & Warmke, 2022) and psychologists (Witvliet & McCullough, 2007; Worthington, 2019) consider forgiveness as one of the different ways of ceasing to blame. Therefore, we also test whether forgiveness entails *ceasing to blame*. Finally, we also consider expectations of *reparations* and an *apology*, as well as a tendency to be *less friendly*

⁴ However, they might be exactly the most fitting emotions in more severe cases, such as the ones mentioned in the introduction. Also, we do not claim that even in mundane cases, it would be inadequate for a victim to feel hatred. We only suggest that for an experimental investigation of rather mundane cases, they might not be the most prototypical emotions.

towards the wrong-doer plausible reactions that change after forgiveness is granted.

Study 1: Granting Forgiveness

To investigate the effect of forgiveness on emotions, cognitive states, conative attitudes, and the act of blaming, we use a vignette study with a pre-/post-intervention design. The experimental design and all predictions were [preregistered](#) with the Open Science Framework. All materials are available in this [repository](#).

Methods

We implemented a $2 \times 4 \times 3$ two-way, mixed design, with a pre- and post-intervention structure. Participants were randomly assigned one of the four cover stories describing a morally relevant interaction between a victim and a wrong-doer. They then rated ten statements on a 7-point Likert scale from 1 (definitely not the case) to 7 (definitely the case). The ten statements read as follows, and Table 1 summarises the categorization of these statements:

(Wrongness) Tom thinks that Andrew’s constant interrupting was wrong.

(Blameworthiness) Tom believes that Andrew deserves blame.

(At Fault) Tom thinks that it is Andrew’s fault that Tom could not make a better impression on the clients.

(Anger) Tom is angry with Andrew.

(Sadness) Tom feels sad about Andrew’s behavior.

(Disappointment) Tom is disappointed in Andrew.

(Apology) Tom expects an apology from Andrew.

(Unfriendliness) Tom will be less friendly in his future interactions with Andrew.

(Reparation) Tom hopes that Andrew will make it up to him somehow.

(Blame) Tom blames Andrew.

Table 1: Categorization of the ten statements

Category	Statements
Cognitive	Wrongness, Blameworthiness, At Fault
Emotion	Anger, Sadness, Disappointment
Conative	Apology, Unfriendliness, Reparation
Act	Blame

After rating the statements for the first time and providing their rating, participants read the end of the story which contained an apology from the wrong-doer and one of three kinds of responses from the victim:

(Forgiveness Granted) “Thank you for your apology. I forgive you.”

(Forgiveness Denied) “Thank you for your apology, but I do not forgive you.”

(Forgiveness Neither Granted Nor Denied) No response

Participants were asked to rate the same ten statements again. We tested the following four hypotheses.

H1: In the forgiveness granted condition, the mean ratings of the emotional states post-intervention will significantly decrease from pre-intervention.

H2: In the forgiveness granted condition, the mean ratings of the conative states post-intervention will significantly decrease from pre-intervention.

H3: In the forgiveness granted condition, the mean ratings of the cognitive states post-intervention will be significantly above the neutral mid-point (4).

H4: In the forgiveness granted condition, the mean rating of the blame statement post-intervention will be significantly above the neutral mid-point (4).

Participants

450 participants were recruited via Prolific Academics (gender-balanced sample, mean age was 41 years) and completed an online survey on Qualtrics (~38 per condition). Prolific’s internal filters were used to ensure that all participants were at least 18 years old, English native speakers, and had an approval rate of over 80%. They were located in Australia, Canada, New Zealand, the United Kingdom, and the United States of America.

Results

A 2 (pre-/post-) $\times 4$ (cover stories) $\times 3$ (response kind) mixed ANOVA revealed a significant interaction between the within-subject factor, response kind and cover stories, $F(60, 2257.96) = 1.47, p < 0.01, \text{Wilks' Lambda} = 0.818, \text{partial } \eta^2 = 0.03$. The within-subject factor has a significant main effect on the mean ratings, $F(10, 430) = 65.53, p < 0.01, \text{Wilks' Lambda} = 0.395, \text{partial } \eta^2 = 0.604$.

To test **H1** and **H2**, Bonferroni-corrected comparisons were carried out with all the cover stories condition grouped and examining the mean ratings in the forgiveness granted condition, for all ten statements we found a significant decrease from pre-intervention to post-intervention (see Table 2). Therefore, the statements in the emotional category and conative category have significantly decreased from pre-intervention to post-intervention in the forgiveness granted condition, supporting **H1** and **H2**⁵.

condition, 86.2% of them decreased their ratings in the post-condition. Among 84.6 of the participants who rated conative states on average higher than 4 in the pre-condition, 83.3% of them decreased their ratings in the post-condition.

⁵ Exploratory analyses found that, in the forgiveness granted condition, among 97.3% of participants who rated cognitive states on average higher than 4 in the pre-condition, 65.5% of them decreased their ratings in the post-condition. Among 92.6% of participants who rated emotions on average higher than 4 in the pre-

Table 2: The Mean Differences of the ratings between Pre- and Post-intervention in the forgiveness granted condition.

Category	Items	Mean Δ	SE	<i>p</i>
Cognitive	Wrongness	0.52	0.11	<0.001
	Blameworthy	0.84	0.13	<0.001
	At Fault	0.41	0.10	<0.001
Emotion	Anger	1.80	0.14	<0.001
	Sadness	1.00	0.11	<0.001
	Disappointment	1.54	0.14	<0.001
Conative	Apology	3.14	0.17	<0.001
	Unfriendliness	1.17	0.12	<0.001
	Reparation	0.47	0.14	<0.001
Act	Blame	1.11	0.1	<0.001

To test **H3** and **H4**, one-sample t-tests against the neutral mid-point (4) demonstrated that in the post-intervention of the forgiveness granted condition, Wrongness ($M = 5.62$, $SD = 1.40$, $t(148) = 14.08$, $p < 0.001$, Cohen's $d = 1.15$), Blameworthiness ($M = 4.63$, $SD = 1.55$, $t(148) = 4.96$, $p < 0.001$, Cohen's $d = 0.41$), At Fault ($M = 5.68$, $SD = 1.37$, $t(148) = 15.05$, $p < 0.001$, Cohen's $d = 1.23$) and the Act of Blame ($M = 4.77$, $SD = 1.56$, $t(148) = 5.98$, $p < 0.001$, Cohen's $d = 0.49$) were significantly above the neutral mid-point (4). Our result supports **H3** and **H4** (see Figure 1).

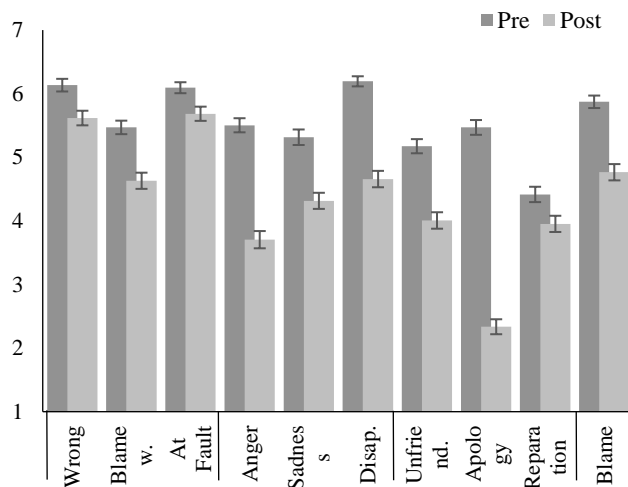


Figure 1: Mean inference ratings per statement. Error bars indicate the standard error around the mean.

Beyond our preregistered hypotheses, we also looked at the items in the forgiveness denied and no response (Bonferroni-corrected). In the forgiveness denied condition, all items except Reparation have significantly increased (the negative mean differences in Table 3). In the no response condition, all items except Reparation have significantly decreased, similar to the forgiveness granted condition.

Table 3: The Mean Differences of the ratings (pre-/post-) in the forgiveness denied and no response conditions.

Cat.	Items	Resp.	Mean Δ	SE	<i>p</i>
Cogn.	Wrong	Denied	-0.31	0.09	<0.001
		No resp.	0.33	0.09	<0.001

Blamew.	Denied		-0.61	0.10	<0.001
		No resp.	0.28	0.10	0.007
At Fault	Denied		-0.37	0.09	<0.001
		No resp.	0.47	0.09	<0.001
Emo.	Anger	Denied	-0.66	0.12	<0.001
		No resp.	1.12	0.11	<0.001
Sadness	Denied		-0.45	0.11	<0.001
		No resp.	0.42	0.11	<0.001
Disap.	Denied		-0.24	0.11	0.025
		No resp.	0.89	0.11	<0.001
Cona.	Apology	Denied	1.07	0.16	<0.001
		No resp.	2.49	0.16	<0.001
Unfriend.	Denied		-0.77	0.10	<0.001
		No resp.	0.88	0.10	<0.001
Repar.	Denied		0.04	0.13	0.755
		No resp.	0.03	0.13	0.804
Act	Blame	Denied	-0.60	0.11	<0.001
		No resp.	0.51	0.10	<0.001

Discussion

Our results provide evidence that cognitive states (Wrongness, Blameworthiness, At Fault), emotional states (Anger, Sadness, Disappointment), conative states (Apology, Unfriendliness, Reparation) and the blame have significantly decreased from pre-intervention to post-intervention, in line with **H1** and **H2**. However, the study also shows that only anger is sufficiently (but not completely) overcome, but not sadness and disappointment, as participants still thought the victim is sad and disappointed after forgiveness. The result also demonstrates the negative emotions impacted by forgiveness include not only anger but sadness and disappointment, although it seems that they are not impacted in the same way.

In line with **H3** and **H4** our results also show that after forgiveness is granted, the mean ratings of the cognitive states and the act of blame are still significantly higher than the mid-point, agreeing with the idea that the victim's cognitive evaluation of the action as wrong and blameworthy is still negative after forgiveness. Interestingly, our study also shows that even though it stays negative, there is a significant decrease in cognitive evaluation, which is currently not accounted for in the literature. Forgiveness likely involves essential changes in the victim's cognitive belief. Another interesting point is that our result does not fully agree with the idea that forgiveness is ceasing to blame: even with a significant decrease, the victim still blames the wrong-doer. This phenomenon will be the focus of our following study.

Study 2: Forgiveness and Blame

Philosophers have argued that forgiveness means overcoming or ceasing to blame. As a consequence, it is often argued that forgiveness is *incompatible* with still blaming the wrong-doer. To investigate the relationship between forgiveness and blame, we use the cancellability test for conversational implicatures (for a theoretical introduction and discussion, see, e.g., Grice, 1989; Sullivan, 2017; Zakkou, 2018; for its empirical application, see, e.g.,

Almeida, Struchiner, & Hannikainen (2021); Baumgartner et al., 2022, Coninx et al., 2022; Willemsen & Reuter, 2021).

In our first study, we apply this experimental paradigm to the concept of forgiveness to determine whether granting forgiveness is indeed considered incompatible with still blaming the wrong-doer.

The experimental design as well as all predictions were [preregistered](#) with the Open Science Framework. All materials are available in this [repository](#).

We implemented a 1×4 one-way, between-subject design. A cancellability statement (hereafter: C-statement) is a two-part conjunction consisting of a combination of forgiveness and blaming. Either Mary has forgiven or not, and the same goes for blaming. This results in four C-statements: Forgiveness and Blame [F&B], No Blame and No Forgiveness [NB&NF], Forgiveness and No Blame [F&NB], and No Blame and Forgiveness [NB&F].

Participants were asked to rate whether the statement was contradictory on a 9-point Likert scale ranging from 1 = “definitely not” to 9 = “definitely yes”. We test the following four hypotheses:

H1: Participants provide significantly higher contradiction ratings for [F&B] than for the Control statements.

H2: In [F&B], participants provide contradiction ratings significantly above the midpoint of the scale.

As previously stated, philosophers and psychologists generally see forgiveness as one way in which we cease to blame. According to this view, ceasing to blame is necessary, yet not sufficient for forgiveness – we can cease to blame for other reasons. Hypotheses 3 & 4 follow from this:

H3: Participants provide significantly lower contradiction ratings for [NB&NF] than for [F&B].

H4: In [NB&NF], participants provide contradiction ratings significantly below the midpoint of the scale.

Participants

238 participants were recruited via Prolific Academics (gender-balanced sample) and completed an online survey on Qualtrics (~60 per condition). Prolific’s internal filters were used to ensure that all participants were at least 18 years old, English native speakers, and had an approval rate of over 80%. They were located in Australia, Canada, New Zealand, the United Kingdom, and the United States of America.

Following previous studies (Baumgartner et al., 2022, Coninx et al., 2022; Willemsen & Reuter, 2021), participants received a short training on the concept of contradiction before entering the actual experiment. Only those participants who correctly answered the test question after this training were included in the analysis (see pre-registration). 6 participants were excluded for failing this screening question.

⁶ Each one-sample t-test here compares the mean of the contradictory rating of a sample of one condition against a constant. Given the between-subjects design, each sample of one condition is

Therefore, all results are presented based on 232 participants, with a mean age of 41.83 ($SD = 14.38$).

Results

The results of Study 2 are depicted in Figure 2. A one-way ANOVA found that there is a significant difference between [F&B], [NB&NF], and the Control, $F(2, 229) = 72.76, p < 0.001$, partial $\eta^2 = 0.39$. A Bonferroni post-hoc test showed that the mean contradiction ratings of both experimental conditions, [F&B] ($5.72 \pm 3.00, p < 0.001$) and [NB&NF] ($5.90 \pm 3.08, p < 0.001$), are significantly higher than that of the Control (1.84 ± 1.84). The two experimental conditions do not differ significantly from one another ($p > 0.99$).

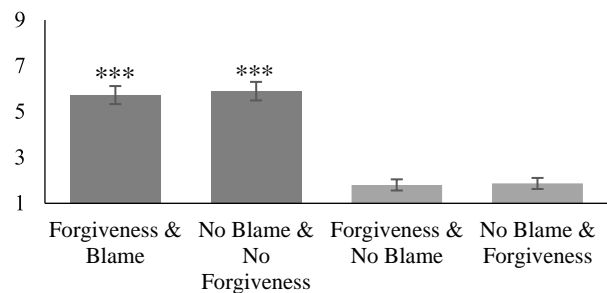


Figure 2: Mean contradiction ratings per statement. Error bars indicate the standard error around the mean.

One sample t-tests against the neutral mid-point (5) reveals that the mean contradiction rating of our second experimental condition [NB&NF] is significantly above the mid-point, $t(57) = 2.22, p = 0.03$, Cohen’s $d = 0.29$.⁶ The mean contradiction rating of our first experimental condition [F&B] is not significantly different from the mid-point, $t(57) = 1.84, p = 0.071$, Cohen’s $d = 0.24$.

Discussion

In line with our prediction **H1**, participants considered a statement in which a speaker claims to have forgiven the wrong-doer but still maintains that she blames her for it [F&B] more contradictory than our control sentences. Scholars who claim that forgiveness is essentially about overcoming blame might see this as direct support of their position. However, participants considered [F&B] statement not *clearly* contradictory as contradiction ratings did not differ from the neutral midpoint. If forgiveness and still blaming were indeed incompatible, higher contradiction ratings should be expected. Therefore, our results do not support **H2**.

While there is a large consensus among theorists that ceasing to blame is *necessary* for forgiveness, none of them claims that ceasing to blame is *sufficient* for forgiveness. Interestingly, when a speaker claims to have ceased to blame yet not forgiven, participants give contradiction ratings above the neural midpoint, and those ratings are not significantly

considered independent and separate from other samples of other conditions.

lower than the ratings of [F&B]. Therefore, our results do not provide evidence for H3 and H4.

General Discussion

Summary of the key findings

Forgiveness plays a significant role in various aspects of daily life, including religion, morality, and even the law. Moreover, it has also been of considerable interest in various academic disciplines, most notably philosophy and psychology. Despite its relevance, we still lack a proper, empirically grounded characterization of the *concept of forgiveness*, and we further lack a proper understanding of *the social practice of granting forgiveness*. In this paper, we empirically investigated what a speaker is understood to convey when they perform the speech act of granting forgiveness by saying, “*I forgive you.*”

Since Emotion Accounts of forgiveness are not only widespread and influential but also intuitively convincing, the first aim of our research was to examine which negative emotions are overcome when a speaker forgives, and whether these negative emotions are overcome completely or only in part. Our results indicate that three emotions that have received considerable support in the literature, namely *sadness*, *anger*, and *disappointment*, played a role in all four scenarios we presented to participants. After the victim expressed their forgiveness, participants judged that the victim felt less of these negative emotions. We did not find that participants believed that these negative emotions are entirely overcome, they are merely felt to a significantly lesser degree (Study 1).

These results are illuminating in two ways. First, our study provides a first answer to the question of which emotions are involved in forgiveness. Sadness, anger, and disappointment seem empirically adequate suggestions. However, the actual list might be significantly longer and may further include hatred or loathing as well—especially in scenarios in which more serious harm or damage is caused. Future studies might manipulate the scenarios to cover a wider range of moral and legal transgressions and more extreme cases, including grave physical or emotional harm or death. Second, in concrete social situations in which a victim grants forgiveness, participants believe that the victim has overcome those negative emotions, yet not entirely. To what extent these insights into the social practice speak to the underlying concept of forgiveness is yet to be discussed.

Blame and Forgiveness

A further aim of our study was to examine the *role of blame* in forgiveness. Philosophers, psychologists, and legal scholars alike have argued that when we forgive, we cease to blame the wrong-doer. Our evidence supports this idea, but only to a certain degree. Participants indeed judge the speaker less likely to blame the agent after they have forgiven

(Study 1), indicating that the agent, at least partly, overcame their initial blame. However, this finding does not support the stronger idea that forgiveness and blame are *incompatible*, as participants believe the agent still blames the wrong-doer. In Study 2, we also find that participants do not consider it contradictory to say, “I have forgiven Hannah for what she did to me and I still blame Hannah for it”.⁷

Drawing attention to the cognitive dimension of forgiveness, we have also investigated some of the beliefs that the victim holds before and after she has forgiven. Our results suggest that a victim does not necessarily overcome the beliefs that the wrong-doer acted wrongly, is blameworthy for the action, and is at fault for the negative consequences that resulted from it. Interestingly, all of those negatively valenced beliefs have decreased significantly, a phenomenon that is currently not accounted for and deserved further empirical investigation.

Our research shows some surprising results. In Study 2, we did not expect that participants consider the statement about a victim who does not blame and does not forgive the agent for the wrongdoing [NF&NB] as contradictory. Philosophers and psychologists regard forgiveness as one way to cease blaming. Consequently, it is conceivable for a person to cease blaming without forgiveness: she might have forgotten, excused, or condoned the agent. However, our result is contrary to this line of thinking. One explanation of this effect is that the folk concept of forgiveness differs from the concept used by philosophers and psychologists. For laypeople, the concept of forgiveness may have a blurry, rather than distinct, boundary to other similar ways of ceasing to blame. Alternatively, one might think that laypeople understand the phrase, “I have not forgiven (the agent)”, as an expression of anger and resentment, and as a signal that an apology or reparations are expected. Further research on the implications of the statement on not forgiving can shed light on this contradiction.

Our studies explicitly followed Haber’s suggestion to investigate the speech act performed by the utterance, “I forgive you.” However, it seems that this explicit mention of forgiveness rarely happens in everyday life and is only one of many ways in which forgiveness can be granted. We usually express forgiveness without mentioning the word ‘forgive’, such as “It’s okay”, or “It’s fine.” It remains to be seen if these ways of forgiveness function similarly or differently.

Study 1 shows that forgiveness reduces the victim’s anger, sadness, and disappointment. However, there are many aspects of forgiveness in relation to emotions that can be investigated: How do the victim’s psychological make-up and the scenarios interact with the emotions involved? What negative emotions, if any, are always overcome when we forgive? Future research on these questions will provide insights into the social practice of forgiving and its underlying psychology.

thought of the sincerity of the statements. Further studies are needed to explore the complex relationship between forgiveness and blame.

⁷ However, there is a possibility that the person uttering such a statement might be insincere in her forgiveness. One limitation of our study is that our design does not access what participants

Acknowledgements

Walker Siu Hong Au and Pascale Willemsen were funded by the Swiss National Science Foundation (SNSF), grant number PZ00P1_201737.

We would like to express our gratitude to Glenn Anderau, Lucien Baumgartner, Christoph Halbig, Ethan Landes, Per-Erik Milam, August Olsen, Kevin Reuter, Stefan Riedener as well as the participants of the Zurich XPhi Lab Meeting for their helpful feedback.

References

- Allais, L. (2008). Wiping the slate clean: The heart of forgiveness. *Philosophy & Public Affairs*, 36(1), 33–68. <https://doi.org/10.1111/j.1088-4963.2008.00123.x>
- Allais, L. (2013). Elective forgiveness. *International Journal of Philosophical Studies*, 21(5), 637–653. <https://doi.org/10.1080/09672559.2013.767525>
- Allen, M. (2015). *Locating forgiveness in criminal law and punishment* [Phd, Birkbeck, University of London]. <http://vufind.lib.bbk.ac.uk/vufind/Record/508235>
- Almeida, G., Struchiner, N., & Hannikainen, I. (2021). Rule is a dual character concept. Available at *SSRN 4018823*.
- Austin, J. L. (1975). *How to do things with words*. Oxford University Press.
- Baumgartner, L., Willemsen, P., & Reuter, K. (2022). The polarity effect of evaluative language. *Philosophical Psychology*.
- Bennett, C. (2018). The Alteration thesis: Forgiveness as a normative power. *Philosophy & Public Affairs*, 46(2), 207–233. <https://doi.org/10.1111/papa.12117>
- Bibas, S. (2006). Forgiveness in criminal procedure. *Ohio State Journal of Criminal Law*, 4, 329.
- Boon, S. D., & Sulsky, L. M. (1997). Attributions of blame and forgiveness in romantic relationships: A policy-capturing study - ProQuest. *Journal of Social Behavior and Personality*, 12(1). <https://www.proquest.com/openview/8f4f6c1ec36113b1d97ca46016cac832/1?pq-origsite=gscholar&cbl=1819046>
- Brunning, L., & Milam, P.-E. (2022). Letting go of blame. *Philosophy and Phenomenological Research*, n/a(n/a). <https://doi.org/10.1111/phpr.12899>
- Coninx, S., Willemsen, P., & Reuter, K. (2022). An experimental-linguistic study of the folk concept of pain: Implication, projection, & deniability. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Darby, B. W., & Schlenker, B. R. (1982). Children's reactions to apologies. *Journal of Personality and Social Psychology*, 43(4), 742–753. <https://doi.org/10.1037/0022-3514.43.4.742>
- Darwall, S. (2006). *The Second-person standpoint: Morality, respect, and accountability*. Harvard University Press.
- Emerson, J. (1964). *The dynamics of forgiveness*. Westminster Press.
- Enright, R. D., & Fitzgibbons, R. P. (2015). *Forgiveness therapy: An empirical guide for resolving anger and restoring hope* (pp. xi, 358). American Psychological Association. <https://doi.org/10.1037/14526-000>
- Fricker, M. (2019). Forgiveness—An ordered pluralism. *Australasian Philosophical Review*, 3(3), 241–260. <https://doi.org/10.1080/24740500.2020.1859230>
- Gehm, J. R. (1992). The function of forgiveness in the criminal justice system. In H. Messmer & H.-U. Otto (Eds.), *Restorative Justice on Trial: Pitfalls and Potentials of Victim-Offender Mediation—International Research Perspectives—* (pp. 541–550). Springer Netherlands. https://doi.org/10.1007/978-94-015-8064-9_38
- Haber, J. G. (1991). *Forgiveness*. Rowman & Littlefield Publishers.
- Heider, F. (1958). *The psychology of interpersonal relations*. Martino Publishing.
- Helmreich, J. S. (2022). Accepting forgiveness. *The Journal of Ethics*, 26(1), 1–25. <https://doi.org/10.1007/s10892-020-09352-0>
- Hieronymi, P. (2001). Articulating an uncompromising forgiveness. *Philosophy and Phenomenological Research*, 62(3), 529–555. <https://doi.org/10.1111/j.1933-1592.2001.tb00073.x>
- Hughes, P. M. (1993). What is involved in forgiving? *The Journal of Value Inquiry*, 27(3), 331–340. <https://doi.org/10.1007/BF01087682>
- Hughes, P. M., & Warmke, B. (2022). Forgiveness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2022). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2022/entries/forgiveness/>
- Lacey, N., & Pickard, H. (2015). To Blame or to forgive? Reconciling punishment and forgiveness in criminal justice. *Oxford Journal of Legal Studies*, 35(4), 665–696. <https://doi.org/10.1093/ojls/gqv012>
- Lerman, D. M. (1999). Forgiveness in the criminal justice system: If it belongs, then why is it so hard to find. *Fordham Urban Law Journal*, 27, 1663.
- Litwinski, L. (1945). Hatred and forgetting. *The Journal of General Psychology*, 33(1), 85–109. <https://doi.org/10.1080/00221309.1945.10544496>
- McCullough, M. E., Fincham, F. D., & Tsang, J.-A. (2003). Forgiveness, forbearance, and time: The temporal unfolding of transgression-related interpersonal motivations. *Journal of Personality and Social Psychology*, 84(3), 540–557. <https://doi.org/10.1037/0022-3514.84.3.540>
- McCullough, M. E., Pargament, K. I., & Thoresen, C. E. (Eds.). (2000). *Forgiveness: Theory, research, and practice* (pp. xviii, 334). The Guilford Press.
- McThenia, P. (1999). The Role of forgiveness in capital murder cases. *Capital Defense Journal*, 12, 325.
- Moore, K. D. (1997). *Pardons: Justice, mercy, and the public interest*. Oxford University Press USA.

- Murphy, J. G. (1988). Forgiveness, mercy, and the retributive emotions. *Criminal Justice Ethics*, 7(2), 3–15. <https://doi.org/10.1080/0731129X.1988.9991835>
- Murphy, J. G. (2004). *Getting even: Forgiveness and its limits*. Oxford University Press USA.
- Nelkin, D. K. (2013). Freedom and forgiveness. In I. Haji & J. Caouette (Eds.), *Free will and moral responsibility* (pp. 165–1
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>
- Nygaard, R. L. (1997). On the role of forgiveness in criminal sentencing | Office of justice programs. *Seton Hall Law Review*, 27(3). <https://www.ojp.gov/ncjrs/virtual-library/abstracts/role-forgiveness-criminal-sentencing>
- Piaget, J. (1932). *Das moralische Urteil beim Kinde* (2nd edition). Klett-Cotta im Dt. Taschenbuch-Verl.
- Pettigrove, G. (2004). The forgiveness we speak: The illocutionary force of forgiving. *The Southern Journal of Philosophy*, 42(3), 371–392. <https://doi.org/10.1111/j.2041-6962.2004.tb01938.x>
- Rokeach, M. (1973). *The nature of human values* (pp. x, 438). Free Press.
- Shapland, J. (2016). Forgiveness and restorative justice: Is it necessary? Is it helpful? *Oxford Journal of Law and Religion*, 5(1), 94–112. <https://doi.org/10.1093/ojlr/rwv038>
- Strelan, P., Feather, N. T., & McKee, I. (2008). Justice and forgiveness: Experimental evidence for compatibility. *Journal of Experimental Social Psychology*, 44(6), 1538–1544. <https://doi.org/10.1016/j.jesp.2008.07.014>
- Warmke, B. (2016). The normative significance of forgiveness. *Australasian Journal of Philosophy*, 94(4), 687–703. <https://doi.org/10.1080/00048402.2015.1126850>
- Weiner, B., Graham, S., Peter, O., & Zmuidinas, M. (1991). Public confession and forgiveness. *Journal of Personality*, 59(2), 281–312. <https://doi.org/10.1111/j.1467-6494.1991.tb00777.x>
- Willemsen, P., & Reuter, K. (2021). Separating the evaluative from the descriptive: An empirical study of thick concepts. *Thought. A Journal of Philosophy*.
- Witvliet, C. V. O., & McCullough, M. E. (2007). Forgiveness and health: A review and theoretical exploration of emotion pathways. In *Altruism and health: Perspectives from empirical research* (pp. 259–276). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195182910.003.0017>
- Witvliet, C. V. O., Worthington, E. L., Root, L. M., Sato, A. F., Ludwig, T. E., & Exline, J. J. (2008). Retributive justice, restorative justice, and forgiveness: An experimental psychophysiology analysis. *Journal of Experimental Social Psychology*, 44(1), 10–25. <https://doi.org/10.1016/j.jesp.2007.01.009>
- Worthington, E. L. (Ed.). (2005). *Handbook of forgiveness*. Routledge. <https://doi.org/10.4324/9780203955673>
- Worthington, E. L. (2019). Definitions, theories, and processes. In *Handbook of forgiveness* (2nd ed.). Routledge.
- Worthington, E. L., & Sandage, S. J. (2016). *Forgiveness and spirituality in psychotherapy: A relational approach* (pp. x, 319). American Psychological Association. <https://doi.org/10.1037/14712-000>
- Zakkou, J. (2018). The cancellability test for conversational implicatures. *Philosophy Compass*, 13(12).
- Zaibert, L. (2009). The paradox of forgiveness. *Journal of Moral Philosophy*, 6(3), 365–393. <https://doi.org/10.1163/174552409X433436>