

The link between Empathy and Forgiveness: Replication and extensions Registered Report of McCullough et al. (1997)'s Study 1

Chi Fung Chan

ORCID: 0000-0002-7425-8421

Department of Psychology, University of Hong Kong
tccf312@connect.hku.hk / timothychan312@gmail.com

[^]Gilad Feldman

ORCID: 0000-0003-2812-6599

Department of Psychology, University of Hong Kong
gfeldman@hku.hk / giladfel@gmail.com

[^]Corresponding author

*Recommended for publication by
Peer Community in Registered Reports
on August 7, 2023*

See recommendation and open peer-review on:

<https://doi.org/10.24072/pci.rr.100444>

In press at Cognition and Emotion
Accepted for publication on July 19, 2024

Author bios:

Chi Fung Chan was a thesis student at the University of Hong Kong during the academic year 2022-3.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgment and decision-making.

Declaration of Conflict of Interest:

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial disclosure/funding:

The project is supported by the University of Hong Kong Teaching Development Grant.

Authorship declaration:

Chi-Fung CHAN conducted the replication as part of his thesis in psychology.

Gilad Feldman guided the project, supervised each step in the project, ran data collection, and edited the manuscript for submission.

Corresponding author

Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR;
gfeldman@hku.hk ; 0000-0003-2812-6599

Rights:

CC BY or equivalent license is applied to the AAM arising from this submission. ([clarification](#))

Citation of the target research article:

McCullough, M. E., Worthington, E. L., & Rachal, K. C. (1997). Interpersonal Forgiving in Close Relationships. *Journal of Personality and Social Psychology*, 73(2), 321–336.
<https://doi.org/10.1037/0022-3514.73.2.321>

Contributor Roles Taxonomy

Role	Chi-Fung Chan	Gilad Feldman
Conceptualization	X	X
Pre-registration	X	X
Data curation		X
Formal analysis	X	
Funding acquisition		X
Investigation	X	
Pre-registration peer review / verification		X
Data analysis peer review / verification		X
Methodology	X	
Project administration		X
Resources		X
Software	X	
Supervision		X
Validation		X
Visualization	X	
Writing-original draft	X	
Writing-review and editing		X

Abstract

McCullough et al. (1997) demonstrated that in situations of feeling hurt by an offender, empathy towards the offender is positively associated with forgiving the offender, which in turn is positively associated with conciliatory behavior and negatively associated with avoidance behavior. In a Replication Registered Report with a Prolific US online sample ($N = 794$), we conducted a replication of Study 1 from McCullough et al. (1997) with extensions manipulating empathy to determine causality and measuring revenge motivation adopted from McCullough et al. (1998). We found that empathy was positively associated with perceived apology ($r=0.45[0.35,0.55]$) and forgiveness toward the offender ($r=0.64[0.56,0.70]$), and forgiveness was positively associated with conciliatory motivation ($r=0.51[0.41,0.59]$) and negatively associated with avoidance motivation ($r=-0.51[-0.59,-0.42]$) and revenge motivation ($r=-0.43[-0.52,-0.33]$). Manipulating empathy, we found that participants who recalled situations in which they felt strong empathy towards the offender rated higher forgiveness compared to participants recalling situations with low empathy or compared to control ($d=0.60-0.62$). Overall, this was a successful replication of the findings by McCullough et al. (1997; 1998) with the empathy model of forgiveness receiving strong empirical support. Materials, data, and code are available on: <https://osf.io/fmuv2/>. This Registered Report has been endorsed by Peer Community in Registered Reports: <https://doi.org/10.24072/pci.rr.100444>.

Keywords: forgiveness, empathy, apology, motivational change, relationship, registered report, replication, social psychology

PCIRR-Study Design Table

Question	Hypothesis	Sampling plan	Analysis plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation on given different outcomes	Theory that could be shown wrong by the outcomes
Is empathy associated with perceived apology and forgiveness?	The relationship between apology and forgiving is largely mediated by empathy. [Reframed as: Apology, forgiving, and empathy are correlated. Empathy causally impacts forgiveness and apology (extension)]	The current study aimed to recruit 800 participants, well-powered enough to detect effects much weaker than the smallest effects in the target. See Power analysis section	Pearson correlation, Between-subject ANOVA (Extension), Bootstrapping mediation (Exploratory)	We followed analyses in the original article and extended it to better address the research questions and report of results. We conducted a power analysis of the target's reported effects, and decided on following the sample size of the target's (239), more than 2.5 times of the required sample (94). Sensitivity analysis indicated the ability to detect correlations of $r = 0.21$ in the control condition. We added 2 more conditions for the extension, resulting in an overall sample of 717 (after exclusions) allowing the detection of $f = 0.15$ (95% power, $\alpha = 5\%$, one-tail). Alpha of 5% followed the target's, and high power of 95% is on par and higher than typical replications in PCIRR.	We examine the replicability of McCullough et al. (1997) and support for our suggested extensions.	That apology, forgiveness, and empathy are correlated. And that empathy impacts apology and forgiveness
Is forgiveness associated with behavioral motivations?	Forgiving promotes constructive actions toward the offender (i.e., conciliation) and inhibits destructive actions toward the offender (i.e., avoidance and revenge) following an interpersonal offense		Pearson correlation			Forgiving as a motivational transformation that inclines people to inhibit relationship-destructive responses and to behave constructively toward someone who has behaved destructively toward them

The link between Empathy and Forgiveness:

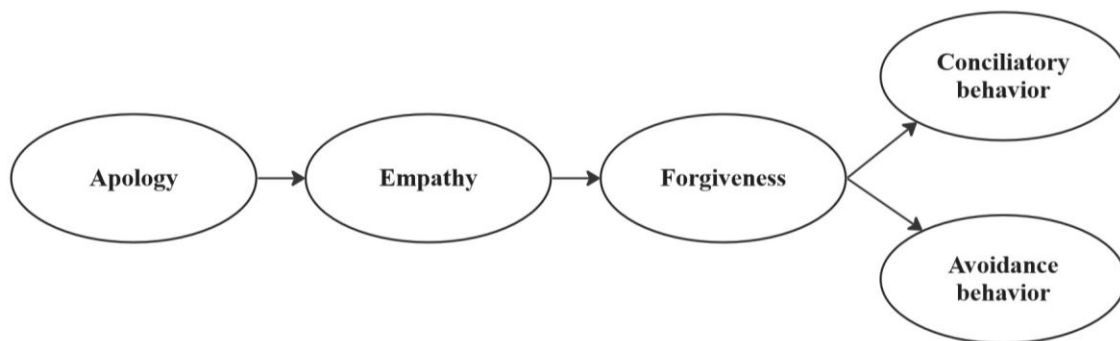
Replication and extensions Registered Report of McCullough et al. (1997)'s Study 1

Background

Many theories and models have been suggested to explain forgiveness and its social roots and implications (e.g., Enright & Coyle, 1998; Strelan & Covic, 2006; Worthington & Scherer, 2004). McCullough et al. (1997)'s empathy model conceptualized forgiveness as the motivation to inhibit relationship-destructive responses and behave constructively toward an offender. Their research demonstrated that (a) relationship between receiving an apology from and forgiving one's offender is a function of increased empathy for the offender and (b) forgiving is uniquely related to conciliatory behavior and avoidance behavior toward the offending partner. Their empathy model of forgiveness is summarized in Figure 1.

Figure 1

Empathy model of forgiveness reconstructed from McCullough et al. (1997)



We report a close replication and extension Registered Report of McCullough et al. (1997) with two main goals. Our first goal was to conduct an independent close replication of the associations among empathy, perceived apology, forgiving, and various behavioral motivations. Our second goal was to extend the target article's design to examine causality by manipulating empathy attributions and incorporating avoidance motivation and revenge motivation measures from a related follow-up study by McCullough et al. (1998) (Study 1). Together, we aimed for a broader, causal, more extensive view on the associations and impact of empathy.

We begin by introducing the literature on forgiveness and the chosen article for replication. We discuss our motivation for the current replication study, the hypotheses, and study design, with our adjustments and added extensions.

Interpersonal forgiveness

Despite different definitions across contexts and theories, forgiveness is generally agreed by scholars to be an intentional and voluntary process (or the result of a process) that involves a change in emotion and attitude regarding an offender, driven by a deliberate decision to forgive (Enright & Fitzgibbons., 2000; Fincham et al., 2004; Worthington & Scherer, 2004). This process usually leads to decreased motivation to retaliate or maintain estrangement from an offender, and requires setting free of negative affects toward the offender (Macaskill, 2012; Webb & Toussaint, 2019).

Field et al. (2013) generalized forgiveness into four major components: Self-awareness, letting go, perspective-taking, and moving on. Whether an *apology* and *reconciliation* are necessary for forgiveness remains controversial among theorists (Fincham et al., 2004; Kelley et al., 2018; Strabbing, 2020). The main arguments against their necessity include the

inapplicability of relationship restorations to self-forgiveness (Hall & Fincham, 2005) and the impossibility of receiving an apology from those who have passed away (Breitbart, 2018; Gassin & Lengel, 2014), and yet forgiveness is generally believed as sensible in both situations. Putting aside this unsolved debate, we mainly focus on McCullough et al. (1997)'s *interpersonal forgiveness* in which apology and reconciliation are possible.

The benefits of interpersonal forgiveness have been widely studied, as - for example - forgiveness seems crucial for psychological healing in broken relationships (Menahem & Love, 2013). Lee and Enright (2019)'s meta-analysis indicated a positive association between forgiveness and physical health (e.g., lowering blood pressure and cortisol levels, improving the immune system), and there is evidence in support of a positive association with marital adjustment (Agu & Nwankwo, 2019; Fahimdanesh et al., 2020; McNulty, 2008), quality of friendships (Boon et al., 2022) and familial relationships (Gordon et al., 2009; Maio et al., 2008).

Empathy in forgiveness

McCullough et al. (1997) conceptualized empathy as a crucial facilitative condition for overcoming the tendency toward destructive responding following an interpersonal offense, leading to forgiveness. Their hypothesis was based on the facilitating effect of empathy occurring in other prosocial phenomena, such as in corporations, altruism, and the inhibition of aggressions. (Batson et al., 1991; Eisenberg & Fabes, 1990; Hoffman, 1981; Moore, 1990; Tangney, 1991; as cited in McCullough et al., 1997). More recent studies also further supported this argument, showing a close association between empathy and forgiveness, across genders (Mellor et al., 2012; Miller et al., 2008), and contributing to personal self-esteem (Turnage et al., 2012; Yao et al., 2017).

Empathy can be treated as either an affect (emotion) in response to stimuli or a dispositional (personality) trait of a person. We followed the target article to focus on the affective dimension of empathy, unless mentioned otherwise.

Revenge, Avoidance, and Conciliation

McCullough et al. (1997) suggested the primacy of the behavioral tendencies toward revenge and avoidance in response to interpersonal offense is motivated by two key affective responses illustrated by Gottman's (1994) research on a close relationship: righteous indignation (e.g., anger, contempt) and hurt-perceived attack (e.g., internal whining, innocent victimhood). This was largely endorsed by recent research on revenge psychology (Jackson et al., 2019; McCullough et al., 2013; Sjöström & Gollwitzer, 2015).

Based on Batson's Empathy-Altruism hypothesis (Batson et al., 1991; Batson & Charles, 2011), McCullough et al. (1997) proposed the psychological similarity between the relationship among empathy, forgiving, and resulting behavioral responses and the sequence of events by which empathy leads to the motivation to care for others (i.e., altruism) and how that altruistic motivation can produce behavioral outcomes (e.g., helping, allocating resources in a social dilemma, cooperating). Therefore, McCullough et al. (1997) suggested empathy may counteract the motivation of relationship-destructive response of revenge and avoidance, in response to interpersonal offense, promoting conciliation through forgiveness.

Choice of study for replication

We chose McCullough et al. (1997)'s study based on two factors: impact, potential for further extensions examining causality and revenge, and the absence of direct replications.

The article has had an impact on scholarly research, especially in the domains of social and clinical psychology. At the time of writing (January 2023), there were 2404 Google Scholar citations of the article with important follow-up theoretical and empirical articles, such as Thompson et al. (2005) on the dispositional dimension of forgiveness and Raes et al. (2011) on the construction of the self-compassion scale. The influential 2-component motivational model of forgiveness proposed by McCullough et al. (1998) was also an extension built on this initial empathy model. The empathy-forgiveness link demonstrated by McCullough et al. (1997)'s research has been one of the most critical foundations of forgiveness therapy which is nowadays widely adopted in clinical settings (Akhtar & Barlow, 2018; Enright & Fitzgibbons, 2015; Yu et al., 2021).

McCullough et al. (1997)'s study is considered one of the first theoretical and empirical bases to explore forgiveness, conceptualizing forgiveness and its correlated factors, providing a new framework to understand forgiveness, transforming it from a sacred virtue or a remote moral standard into an explainable social phenomenon. This has led to further studies of the implications of forgiveness aiming to aid the public in improving social well-being and interpersonal relationships in their daily lives (Akhtar & Barlow, 2018; Worthington et al., 2007). McCullough et al. (1997)'s research offered a scientific framework for forgiveness intervention and psychotherapy. The empathy-forgiving link was the theoretical foundation for several psychological treatments and therapies for a variety of life problems and mental illnesses in clinical settings, ranging from spousal infidelity (Chi et al., 2019) or bereavement (Záhorcová et al., 2021), to borderline personality disorder (Sandage et al., 2015) or post-traumatic stress disorder (Akhtar & Barlow, 2018; Currier et al., 2016).

The target article suggested what appears to be a causal model (see Figure 1), and yet the methods employed to test the mediation were based on correlational designs. We saw potential in extending their design with modifications aiming to establish the causality of the impact of empathy on forgiveness.

To the best of our knowledge, there are currently no published independent direct replications of this article. McCullough et al. (1998) extended their model by adding other variables such as commitment, impact of the offense, and rumination, into predicting forgiveness, which we aimed to further integrate into our replication as an extension. Donovan et al. (2020) conducted a related conceptual replication of the Model of Motivated Interpersonal Forgiveness, with different measurements and designs.

Following the recent growing recognition of the importance of reproducibility and replicability in psychological science (e.g., Brandt et al., 2014; Open Science Collaboration, 2015; Nosek et al., 2022; Zwaan et al., 2018), we aimed to revisit the classic Empathy Model of Forgiveness with a well-powered close independent replication Registered Report of McCullough et al. (1997), integrating developments from McCullough et al. (1998) and aiming to test causality.

Original hypotheses and findings in the target article

McCullough et al. (1997) conceptualized interpersonal forgiving as the set of motivational changes whereby one becomes (a) less motivated to retaliate against an offending relationship partner, (b) less motivated to maintain estrangement from the offender, and (c) more motivated towards conciliation and goodwill for the offender, despite the offender's hurtful actions. Affective empathy was conceptualized as a crucial facilitative condition for overcoming the

primary tendency toward destructive responses following a significant interpersonal offense. On the basis of these conceptual analyses, McCullough et al. (1997) proposed three core hypotheses. We summarized the hypotheses of the target article in Table 1.

Table 1

Summary of hypotheses of the target article

Hypothesis	Description
1	Empathy mediates relationships between dispositional and environmental variables and their causal effects on forgiving.
	a There is a positive association between a wronged person's empathy for an offender and reported forgiveness for the offender.
	b Apology increases the likelihood of forgiving, mediated by empathy.
2	Forgiving promotes constructive actions toward the offender and inhibits destructive actions toward the offender following an interpersonal offense.
	a Forgiveness is positively associated with conciliation motivation.
	b Forgiveness is negatively associated with (i) avoidance motivation and (ii) revenge motivation.
	c Forgiving is causally more proximal (and more strongly related) to behavioral motivation (i.e., conciliation, avoidance and revenge) than is empathy.
3*	Clinical efforts to influence clients' capacity to forgive will succeed insofar as they induce empathy for the offender.

Note. Hypothesis 3 is not included in the replication because it involves a clinical intervention.

We mainly focused on Study 1 of the target article, examining the link between apology, forgiving, and empathy for offending partners and whether forgiving is associated with increased conciliation and decreased avoidance motivation following the offense.

In the target article's study, the authors recruited a sample of university undergraduates who were asked to think of a particular person who treated them unfairly and hurt them at some point in the past. After visualizing and re-experiencing the situation again, participants described

the interpersonal injury they had received and then completed the empathy, forgiving, and behavioral self-report measures. We summarized the associations reported in the target article in Table 2, adopted from the target article.

Table 2

Target article: Means, Standard Deviations, Internal Consistency Reliabilities and Intercorrelations

Variables	<i>M</i>	<i>SD</i>	α	1	2	3	4	5
1. Degree of apology	5.63	2.84	.79	—				
2. Empathy	13.22	5.95	.88	.36**	—			
3. Forgiving	16.82	6.73	.87	.43**	.67**	—		
4. Conciliatory behavior	6.74	2.50	.74	.44**	.63**	.70**	—	
5. Avoidance behavior	10.11	3.89	.90	-.47**	-.58**	-.73**	-.56**	—

Note. Apology scores ranged from 2 to 10. Empathy scores ranged from 0 to 20. Forgiving Scale scores ranged from 5 to 25. Conciliatory behavior scores ranged from 2 to 10. Avoidance behavior scores ranged from 3 to 15. ** $p < .01$. Adopted from McCullough et al. (1997), p. 325

Extension: Examining causal link with empathy manipulation

We aimed to extend the replication study by manipulating empathy. McCullough et al. (1997) indicated that one of the major limitations of their Study 1 was the correlational study design, limiting causal claims implied in their model. We used the target's Study 1 as our control condition, and added two additional conditions manipulating empathy in the recalled situation.

Our main focus was the replication, with the extension added as an exploratory direction. Therefore, in our extension we used the same recall method about the elicited past experience,

and built on top of that. Our aim with the extension was to manipulate the elicitation of recalled situations in which empathy has been experienced so that the person can reflect and evaluate other factors in that situation. Therefore, the manipulation is of the recalled past experience and not the empathy that the participant is experiencing while taking part in the experiment. This is different from some of the research that tried to manipulate empathy through a perspective-taking approach for emotions experienced during the experiment., in which participants were asked to remain objective (vs. emotionally-attached) to the main character when reading a scenario (Berenguer, 2007, 2010).

Our extension approach of manipulating elements of a recalled past event is therefore aligned with the replication and follows commonly used methods in social psychology that study evaluations of emotion laden situations. We previously implemented similar manipulations in recall tasks in various judgment and decision making replication projects (e.g., Chen et al., 2023; Yeung & Feldman, 2022), both based on classic articles in the literature that have previously employed a manipulation of factors in the recalled scenario (e.g., Carter & Gilovich, 2012; Gilovich & Medvec, 1994).

Pre-registration and open-science

We provided all materials, data, and code on: <https://osf.io/fmuv2/>. This project received Peer Community in Registered Report Stage 1 in-principle acceptance (<https://osf.io/q78fs/>; <https://rr.peercommunityin.org/articles/rec?id=380>) after which we created a frozen pre-registration version of the entire Stage 1 packet (<https://osf.io/c7m3v/>) and proceeded to data collection. All measures, manipulations, exclusions conducted for this investigation are reported, and data collection was completed before analyses.

Method

Power analysis

We calculated effect sizes (ES) and power based on the statistics reported in the target article. Both the ES and power were computed using R studio (Version: 1.4.2) with packages “MBESS” and “pwr”. We focused on the intercorrelations between the variables, aiming for a power of 0.95 with an alpha 0.05. The largest minimum sample size required for the correlational tests reported with significant results (i.e., apology vs. empathy) was 94 participants. The calculation was based on the effect size of $r = .36$, with a power of 0.95 and an alpha of 0.05.

To ensure we have enough power to detect all the effects in the target article, we decided that the sample size in our replication should not be lower than the sample size in the target article’s study. Thus, we followed the target article’s sample size of 239 participants. We conducted a sensitivity analysis and found this sample is enough to detect correlations of $r = 0.21$, which is weaker than the lower bound of the weakest effect in the target article (apology vs. empathy: $r = 0.36$, 95% CI [0.24, 0.47]).

In our extension, we added two extra conditions by manipulating empathy, and therefore decided to multiply the sample by three to 717 participants. Accounting for possible exclusions of 0-10% based on our previous experience with the target sample, our integrated design, and allowing for the potential of additional analyses, we aimed for a larger total sample of 800 participants. A sensitivity analysis indicated that a sample of 717 (after exclusions) would allow the detection of $f = 0.15$ for a three-conditions ANOVA for our experimental design (95% power, alpha = 5%, one-tail) for our extension. Also, the sample would be sufficiently powered to detect contrasts of $d = 0.33$ (95% power, alpha = 5%, two-tail), which correspond to a medium effect in

social psychology research (Xiao et al., 2023). Based on our previous experience, recall tasks in judgment and decision-making tended to show medium to very strong effects (e.g., Chen et al., 2023; Feldman et al., 2016; Yeung & Feldman, 2022).

Participants

We recruited a total of 794 US American student participants using Prolific ($M_{age} = 28.8$, $SD = 12.2$). We provided a comparison of the target article sample and our replication and extension sample in Table 3.

We first pretested survey duration with 30 participants to make sure our time run estimate was accurate and adjusted pay as needed, the data of the 30 participants was not analyzed other than to assess survey completion duration and needed pay adjustments. These participants were included in the overall analyses.

Experimental design

We summarized the experimental design in Table 4, a between-subject experimental design with one independent variable and three conditions. We manipulated empathy towards the offender in the recalled situation (i.e., High empathy vs. Low empathy vs. Control) and compared the intercorrelations of the dependent variables (e.g., perceived apology, empathy, and forgiving).

Table 3

Difference and similarities between target article and replication

	McCullough et al. (1997)	US Prolific
Sample size	239	794
Geographic origin	University undergraduate ¹	US American students
Gender	108 males, 131 females	385 males, 381 females, 28 other/did not disclose
Ethnic group	83% White, 14% Black, 3% other	N.A.
Median age (years)	Unreported	24.0
Average age (years)	19	28.8
Standard deviation age (years)	Unreported	12.2
Medium (location)	Unreported	Computer (online)
Compensation	Extra course credit	Nominal payment
Year	1997	2023

Note. ¹ Origin was not explicitly mentioned in the target article, though we suspect it was US American, given the authors' affiliation at the time.

Table 4

Replication and extension experimental design

<u>Level of empathy</u> (between-subject)	<u>High empathy condition</u> (Extension) “you were highly empathetic toward the person who had hurt you.”	<u>Low empathy condition</u> (Extension) “you were not empathetic toward the person who had hurt you”	<u>Control condition</u> (Replication) No indication of empathy towards the offender
Dependent variables (DV)	<u>Offense-related information</u> Questions include: “What was your relationship with the person who had hurt you” “How long has it been since the event occurred” “Please indicate the degree to which the offense had hurt you” (1 = <i>Hurt very little</i> to 5 = <i>Hurt so much</i>) “The person was not wrong in what he/ she did to me.” (0 = <i>Strongly disagree</i> to 5 = <i>Strongly agree</i>) (Source: McCullough et al., 1997)		
	<u>Perceived apology</u> “The offender has apologized?” “The offender has attempted to explain their hurtful behavior?” (1 = <i>Strongly disagree</i> to 5 = <i>Strongly agree</i>) (Source: McCullough et al., 1997)		
	<u>Empathy</u> “Please rate each adjective to indicate the degree to which you feel each of the following affects for the offender” Sympathetic, empathic, concerned, moved, compassionate, warm, softhearted and tender (0 = <i>Not at all</i> to 5 = <i>Extremely</i>) (Source: Batson et al., 1982) [Note: McCullough measured all 8 emotions, though only some analyzed. We kept all 8.]		
	<u>Forgiving</u> “I wish him/her well”, “I disapprove of him/her”, “I think favorably of him/her” and “I condemn the person.” (0 = <i>Not at all</i> to 5 = <i>Extremely</i>) “I have forgiven the person.” (1 = <i>I have not at all forgiven</i> to 5 = <i>I have completely forgiven</i>) (Source: Wade, 1989)		
	<u>Conciliatory motivation</u> “I tried to make amends” “I took steps toward reconciliation: Wrote them, called them, expressed love, showed concern...” (1 = <i>Strongly disagree</i> to 5 = <i>Strongly agree</i>)		

Avoidance motivation

"I keep as much distance between us as possible"

"I live as if he/she doesn't exist, isn't around"

"I don't trust him/her"

"I find it difficult to act warmly toward him/ her."

"I avoid him/her"

"I cut off the relationship with him/her."

"I withdraw from them"

(1 = *Strongly disagree* to 5 = *Strongly agree*)

(Source: McCullough et al., 1998)

Revenge motivation

"I'll make him/her pay"

"I wish that something bad would happen to him/her"

"I want him/her to get what he/she deserves"

"I'm going to get even."

"I want to see him/her hurt and miserable."

(1 = *Strongly disagree* to 5 = *Strongly agree*)

(Source: McCullough et al., 1998)

**Comprehe
nsion
checks
(CC)**

Questions include:

(1) What type of behavior are you asked to recall?

(Someone treated me badly / Someone treated me nicely / Someone had an interaction with me)

(2) How many people are you asked to focus on?

(1 / 2 / 3)

Note. CC questions were newly designed for this replication study but did not exist in the original article.

Procedure

Participants completed the survey via the online survey system Qualtrics. We employed the Qualtrics fraud and spam prevention measures: reCAPTCHA, prevent multiple submissions, prevent ballotstuffing, bot detection, security scan monitor, and relevantID.

Participants were randomly assigned into the experimental conditions: High Empathy condition, Low Empathy condition, control condition, which were counterbalanced using the randomizer “evenly present” function in Qualtrics.

We designed comprehension checks to ensure that participants fully understood their tasks before responding to our study measures. These comprehension check questions were as follows: “What type of behavior done on you are you asked to recall”, “How many people are you asked to focus on”, and “What emotion towards the offender in the situation are you asked to recall”. Participants needed to answer all the questions correctly before proceeding to the next page.

In the High Empathy condition, participants ($n = 264$) were asked to recall a hurting experience that they were “*highly empathetic to the person who had hurt you*”, whereas in the Low Empathy condition, they ($n = 267$) were asked to recall a hurting experience that they were “*not empathetic to the person who had hurt you*”. Participants then described the interpersonal injury they received and explained the reason why they were empathetic/unempathetic towards the offender, completing the self-report measures of perceived apology, empathy, forgiveness, and behaviors.

The control condition was a replication of the target article and closely followed the study’s design. Participants ($n = 263$) were asked to think of a person whom they experienced as

treating them unfairly and hurting them at some point in the past *without* any indication/reference of empathy towards the offender. Then, participants described the interpersonal injury they received, completing the self-report measure as in the first two conditions.

At the end of the experiment, participants answered a number of funneling questions and provided their demographic information. We provided a more comprehensive overview of the survey procedure in “Instructions and experimental material” in the supplementary.

Manipulations

Participants were randomly assigned to one of the three conditions. We expected that participants in the High Empathy condition would rate higher empathy than those in the Low Empathy condition.

We provided additional details of the manipulations, the differences between the three conditions, the experimental design, and complete scales used in the current replication in “Materials and scales used in the replication + extension experiment” in the Supplementary Materials.

Measures

Replication

Offense-related information

Similarly to the target article, participants indicated their age, gender, relationship with the person who had hurt them, the time since the offense occurred, and a brief description of the offense. Many of the offenders whom participants described were romantic partners (29.6%), relatives (18.0%), or friends of the same gender (14.5%).

Then, participants indicated the degree to which the offense hurt them: “The person was not wrong in what he/she did to me” (0 = *Strongly disagree*; 5 = *Strongly agree*).

Perceived degree of apology

We measured the extent to which participants perceived that the offender apologized for the offense with a scale consisting of two items (1 = *Strongly disagree*; 5 = *Strongly agree*; scores ranging from 2 to 10; Cronbach’s alpha (α) = .85, McDonald’s omega (ω) = .85).

Affective empathy

Participants rated the degree to which they felt toward their offender using the Batson’s eight-item empathy scale (Archer et al., 1981; Batson et al., 1986; Batson et al., 1983; Coke et al., 1978; Fultz et al., 1986; Tbi & Batson, 1982), refined by McCullough et al. (1997) into four emotions (empathic, concerned, moved, softhearted) (0 = *Not at all*; 5 = *Extremely*; score ranging from 0 to 20; α = .90, ω = .91).

Forgiveness

We assessed the degree to which the respondent experienced a constructive disposition and the absence of a destructive disposition in light of the offending partner’s hurtful actions using the target’s five-item measure of forgiving : “I wish him/her well”, “I disapprove of him/her”, “I think favorably of him/her”, “I condemn the person” and “I have forgiven the person”. The first four items were on a 6-point scale (0 = *Strongly disagree*; 5 = *Strongly agree*). The final forgiving item was on a 5-point scale (1 = *I have not at all forgiven*; 5 = *I have completely forgiven*). Scores ranged from 1 to 25 (α = 0.86, ω = 0.87).

Conciliatory motivation toward the offender

Two items measured engagement in two reconciliation behaviors: “I tried to make amends” and “I took steps toward reconciliation: Wrote them, called them, expressed love,

showed concern, etc.". (1 = *Strongly disagree*; 5 = *Strongly agree*; scores ranged from 2 to 10; $\alpha = 0.83$, $\omega = 0.83$).

Avoidance and Revenge motivations toward the offender

This measure was derived from the transgression-related interpersonal motivations (TRIM) inventory developed by McCullough et al. (1998). TRIM aimed at evaluating respondents' motivation to avoid personal and psychological contact with the offender (i.e., avoidance) and the motivation to seek revenge or see harm come to the offender (i.e., revenge).

The seven avoidance motivation items were: "I keep as much distance between us as possible", "I live as if he/she doesn't exist, isn't around", "I don't trust him/her.", "I find it difficult to act warmly toward him/her", "I avoid him/her", "I cut off the relationship with him/her", or "I withdraw from him/her" ($\alpha = 0.94$, $\omega = 0.94$). The five revenge motivation items were: "I'll make him/her pay", "I wish that something bad would happen to him/her.", "I want him/her to get what he/she deserves.", "I'm going to get even.", "I want to see him/her hurt and miserable" ($\alpha = 0.89$, $\omega = 0.89$). Both the avoidance motivation and revenge items from TRIM were on a 5-point scale (1 = *Strongly disagree*; 5 = *Strongly agree*). Their overall scores ranged from 7 to 35 and 5 to 25 respectively.

Extension

Offense-related information

Similar to the measures in replication condition detailed above, participants indicated their age, gender, relationship with the person who had hurt them, how much time passed since the offense occurred, as well as the degree of the offense. Yet, in the extension conditions, participants were also asked to explain why they were empathetic/unempathetic to the offender according to their randomly-assigned condition.

Affective empathy

The empathy measure of Batson's eight-item scale in the target article was used as a manipulation check of the empathy manipulation in the extension.

Deviations

We made minor adjustments to the target article in several aspects, summarized in Table 5.

Table 5

Comparison of target article versus replication

	Target article	Replication	Reason for change
Study design	Participants completed the studies with pen and paper in the laboratories.	Participants completed the studies on an online survey.	Lower cost and higher efficiency.
Sample characteristics	<i>N</i> = 239 Sample origin: University undergraduate	<i>N</i> = 794 US American Prolific students	Two extra conditions in extension; Generalizability of results by including a wider variety of participants.
Procedure	Items of all dependent variables (DV) were not randomized Three items for the measure of Avoidance motivation No comprehension check	Items of all DV were randomized Extended to seven items Comprehension checks exist	Addressing potential order effects. McCullough et al. (1998) modified their measure of avoidance motivation with the TRIM inventory Ensuring participants read and understood the task.
Statistical analysis	Pearson's <i>R</i> ; Nested structural equation models (SEM)	Pearson's <i>R</i>	The SEM used in the original article were based on correlations. Our extension changed to testing causality. We toned down the importance of the causal chain, and changed mediation to an exploratory analysis.
Conditions	1 condition	1 conditions identical to the target (Control) with 2 extension conditions.	To examine the causal relationship between empathy and forgiveness in Extension

Evaluation criteria for replication findings

McCullough et al. (1997) conducted 10 correlation tests in the target article as shown in Table 2. We aimed to compare the replication's control condition effects with those in the target article using the criteria set by LeBel et al. (2019) (see subsection "Replication evaluation" in the Supplementary Materials).

Replication closeness evaluation

We provided details on the classification of the replications using the criteria by LeBel et al. (2018) criteria in Table 6 (see section "replication closeness evaluation" in the supplementary). We did not set out to replicate the nested structural equation models (SEMs) in the original article due to its limitation in testing the proposed causal chain (Table 1). LeBel et al. (2018) did not consider statistical tests as an important criterion in its replication closeness evaluation, yet we considered it relevant for replication research. Thus, we summarized the replication as a *close* replication (rather than a *very close* replication).

Table 6

Classification of the replication, based on LeBel et al. (2018)

Design facet	Replication	Details of deviation
Effect/hypothesis	Same	
IV construct	Same	
DV construct	Same	
IV operationalization	Same	
DV operationalization	Similar	The transgression-related interpersonal motivations inventory (TRIM, McCullough et al., 1998) is incorporated to examine the behavioral motivation of avoidance and revenge.
Population (e.g., age)	Similar	Students were recruited through an online research platform Prolific using their demographic filtering.
IV stimuli	Same	
DV stimuli	Same	
Procedural details	Similar	Order of items were randomized
Physical settings	Different	Experiment is conducted online instead of via traditional paper survey
Contextual variables	Different	Participants were recruited online using Prolific.
Replication classification	Close replication	

Data analysis strategy

Replication: Correlation tests

We conducted Pearson's correlations to examine the associations between the six measures of interest: perceived apology, affective empathy, forgiving, conciliatory motivation, avoidance, and revenge motivation.

We did not replicate the full three nested structural equation models (SEM) used in the target article in our study. There are limitations in the target article's attempt to establish a causal mediation relationship using SEM (Rohrer et al., 2022), and requires more careful designs and much larger samples than originally employed. Instead, our extensions aimed to manipulate empathy to test causality directly, and flagged the target's mediation analysis in the control condition as an exploratory analysis, and not core to assessing the target's claims or replicability.

We employed Diedenhofen and Musch (2015)'s "cocor" R package for correlation comparisons. The "cocor" R package is powerful and comprehensive since it compares overlapping correlations from dependent groups with up to 10 commonly used approaches (i.e., Dunn & Clark, 1969; Hendrickson et al., 1959; Hittner et al., 2003; Hotelling, 1940; Meng, 1992; Olkin, 1967; Pearson & Filon, 1898; Steiger, 1980; Williams, 1959; Zou, 2007). These would be conducted as exploratory analyses for addressing Hypothesis 2c.

Extension: the impact of Empathy on Forgiveness

We conducted two between-subjects ANOVAs to examine how apology and forgiveness differ across the three conditions (high empathy vs. low empathy vs. control). We conducted post-hoc tests contrasting condition pairs. We chose post-hoc Scheffe tests because they are the

most conservative post-hoc pairwise comparison method, generating the widest confidence intervals of group means difference.

Outliers and exclusions

In this study, we did not classify outliers. We included all the data collected in our analysis. See section “Exclusion criteria” in the supplementary for more details.

Results

We summarized descriptive statistics in Table 7, correlations in Table 8, and the statistical test results with interpretation in Table 9. We conducted the analyses with R (Version: 4.1.2).

Table 7

Descriptive statistics for all conditions

	High empathy (Extension) (<i>n</i> = 264)	Low empathy (Extension) (<i>n</i> = 267)	Control (Replication) (<i>n</i> = 263)	Overall (<i>N</i> = 794)
Perceived apology	5.18 [2.94]	4.39 [2.7]	4.29 [2.61]	4.62 [2.78]
Empathy	9.56 [5.68]	3.85 [4.72]	4.69 [5.24]	6.03 [5.79]
Forgiveness	14.9 [6.22]	11.1 [6.4]	10.98 [6.38]	10.98 [6.38]
Conciliatory motivation	5.58 [2.68]	4.3 [2.53]	4.52 [2.54]	4.8 [2.64]
Avoidance motivation	20.47 [9.31]	24.15 [9.19]	24.37 [9.2]	23 [9.39]
Revenge motivation	8.01 [4.34]	9.05 [4.82]	9.06 [4.73]	8.71 [4.66]

Note. Format: Mean [standard deviation]. Perceived apology ranged from 2 to 10.

Empathy ranged from 0 to 20.

Forgiveness ranged from 1 to 25.

Conciliatory motivation ranged from 2 to 10.

Avoidance motivation ranged from 7 to 35.

Revenge motivation ranged from 5 to 25.

Table 8*Control condition (Replication): Intercorrelations with confidence intervals*

Variable	<i>M</i>	<i>SD</i>	α	ω	1	2	3	4	5
1 - Perceived apology	4.29	2.61	0.85	0.85					
2 - Empathy	4.69	5.24	0.90	0.91	.45*** [.35, .54] (.43)				
3 - Forgiveness	10.98	6.38	0.86	0.87	.34*** [.23, .44] (.34)	.64*** [.55, .70] (.63)			
4 - Conciliatory motivation	4.52	2.54	0.83	0.83	.26*** [.15, .37] (.29)	.51*** [.41, .59] (.50)	.45*** [.35, .54] (.45)		
5 - Avoidance motivation	24.37	9.2	0.94	0.94	-.28*** [-.40, -.17] (-.27)	-.51*** [-.60, -.42] (-.50)	-.73*** [-.78, -.67] (-.72)	-.32*** [-.42, .20] (-.33)	
6 - Revenge motivation	9.06	4.73	0.89	0.89	.03 [-.09, .15] (.05)	-.11 [-.23, .01] (-.15)	-.43*** [-.52, -.33] (-.45)	-.05 [-.17, .07] (-.04)	.37*** [.26, .47] (.36)

Note. Correlations in the control condition for the replication ($n = 263$). Format: Pearson's correlations [confidence interval] (Spearman's rho).

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table 9*Summary of statistical tests and their interpretation*

Hypothesis	Target article			Replication			Interpretation
	<i>p</i>	Effect size	CI	<i>p</i>	Effect size	CI	
1a	< .001	$r = .67$	[0.59, 0.73]	< .001	$r = .63$	[0.56, 0.70]	Signal, consistent
1b	< .001	$r = .36$	[0.24, 0.47]	< .001	$r = .45$	[0.35, 0.55]	Signal, consistent
2a	< .001	$r = .70$	[0.73, 0.76]	< .001	$r = .45$	[0.34, 0.54]	Signal, inconsistent, smaller
2bi	< .001	$r = -.73$	[-0.78, -0.66]	< .001	$r = -.73$	[-0.78, -0.67]	Signal, consistent
2bii	/	/	/	< .001	$r = -.43$	[-0.52, 0.33]	Signal

Note. Effects are Pearson's correlations. CI = 95% confidence intervals. The interpretations of the outcomes are based on LeBel et al. (2019).

Replication

We conducted 15 Pearson's correlation tests to examine the associations between variables in the control (replication) condition, summarized in Table 8.

First, we found support for hypotheses 1a and 1b that empathy is positively associated with perceived apology, $r(261) = 0.45$, 95% CI [0.35, 0.55], $p < .001$, and forgiveness, $r(261) = 0.64$, 95% CI [0.56, 0.70], $p < .001$. We provided the summary scatterplots for the relationship between perceived apology and affective empathy as well as affective empathy and forgiveness in Figures 2 and 3.

We found support for hypotheses 2a and 2b that forgiveness is positively correlated to conciliation motivation, $r(261) = 0.45$, 95% CI [0.35, 0.54], $p < .001$, and negatively correlated to avoidance motivation, $r(261) = -0.73$, 95% CI [-0.78, -0.67], $p < .001$. We also found support for forgiveness as negatively associated with revenge motivation, $r(261) = -0.43$, 95% CI [-0.52, -0.33], $p < .001$. We provided the summary scatterplots in Figures 4, 5, and 6.

Figure 2

Replication (Control) condition: Association between perceived apology and affective empathy

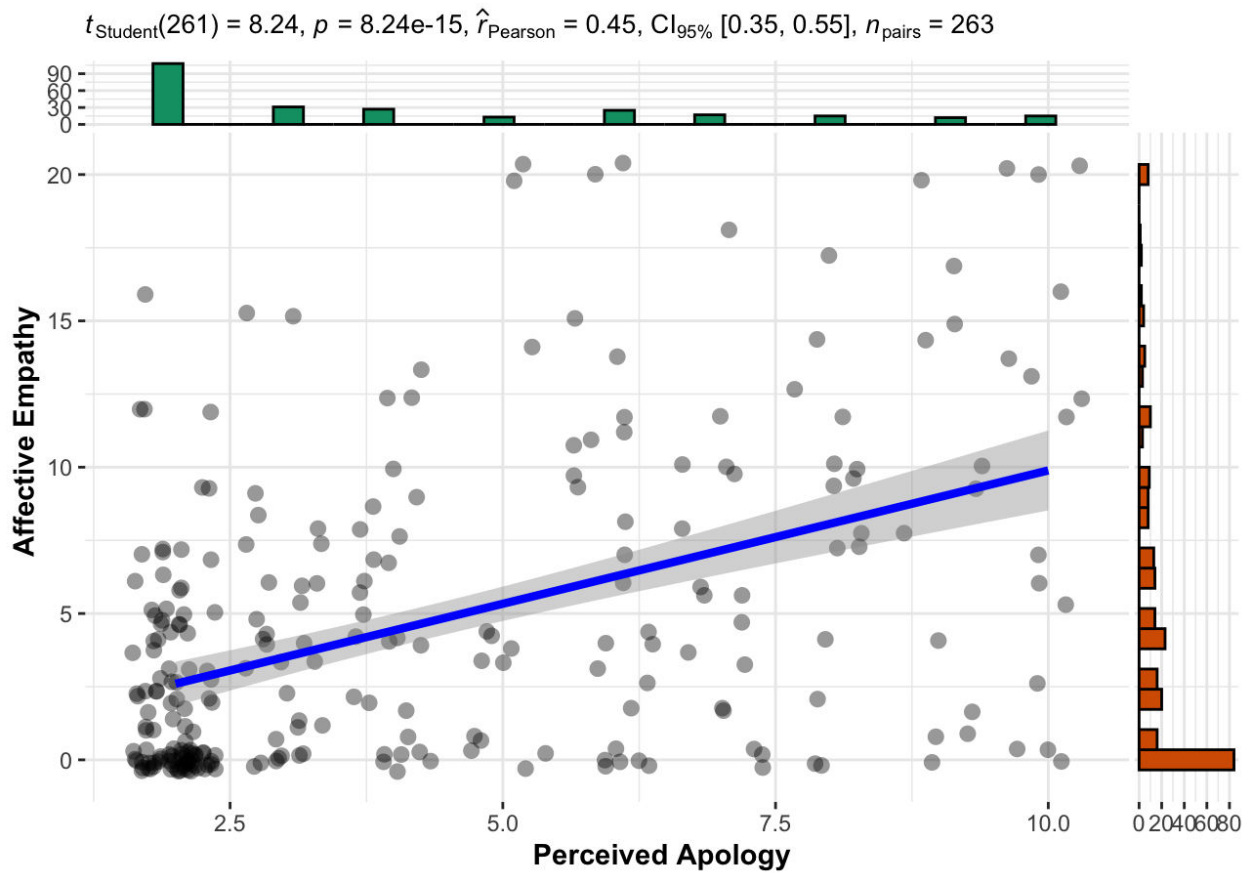


Figure 3

Replication (Control) condition: Association between affective empathy and forgiveness



Figure 4

Replication (Control) condition: Association between forgiveness and conciliatory motivation

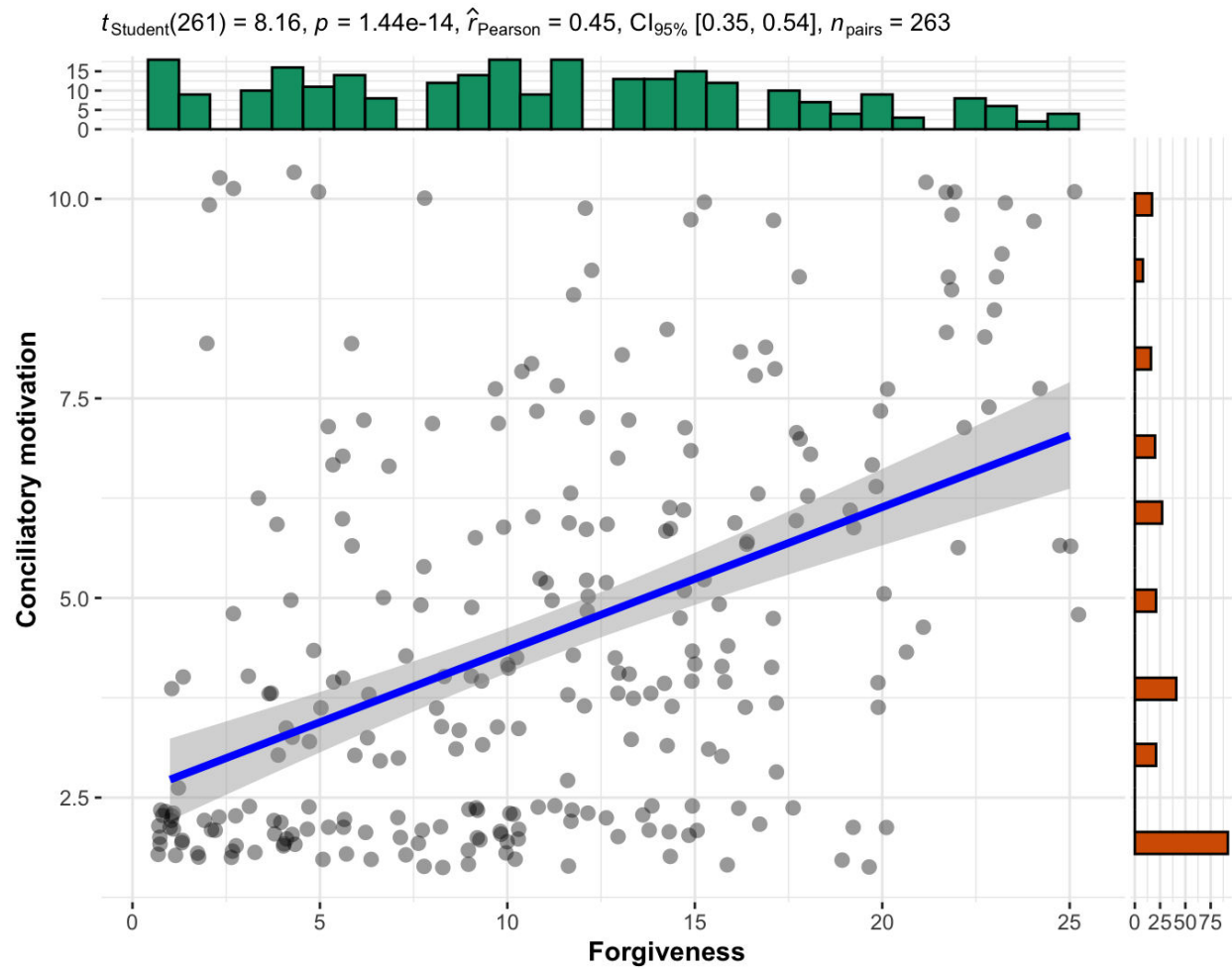


Figure 5

Replication (Control) condition: Association between forgiveness and avoidance motivation

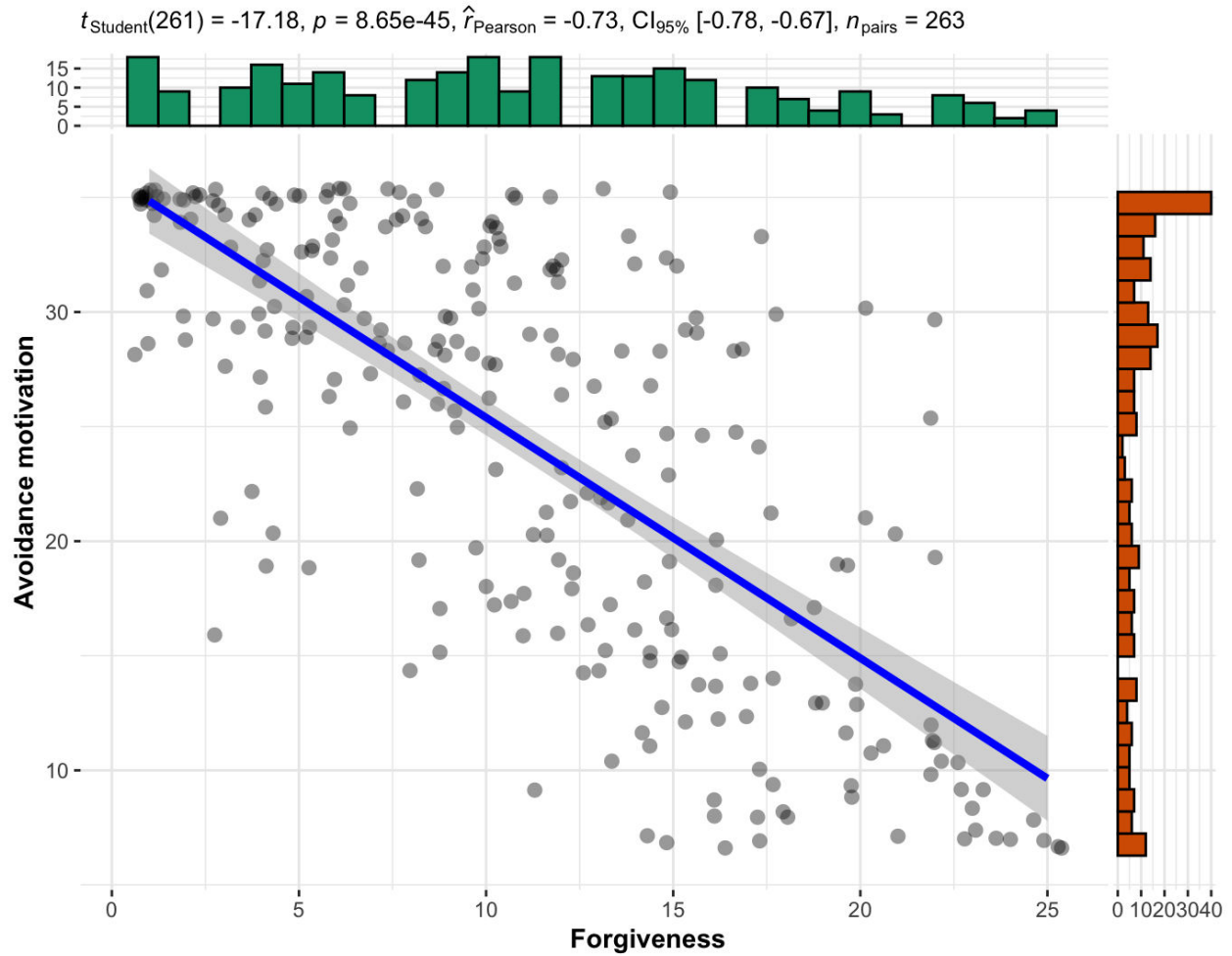
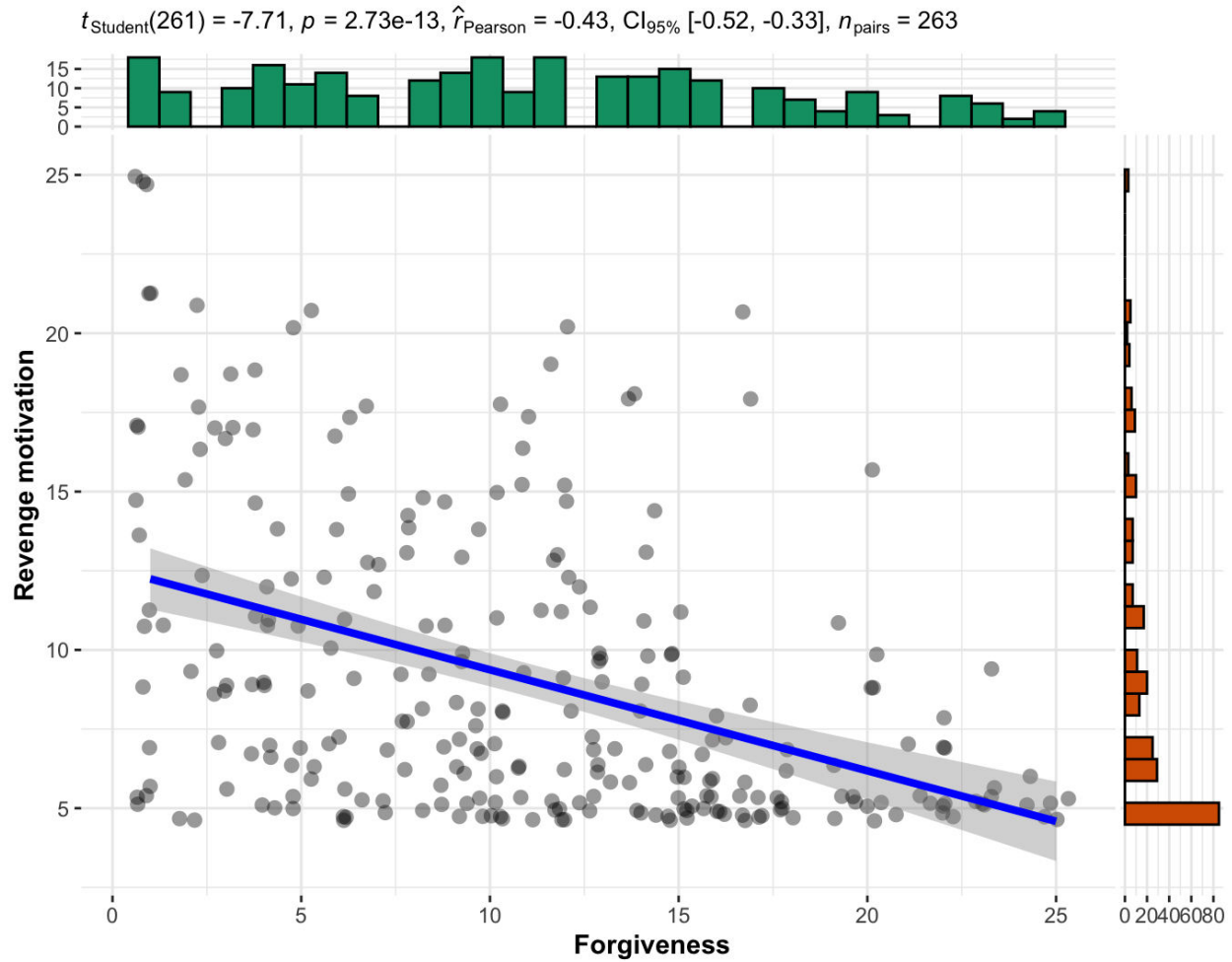


Figure 6

Replication (Control) condition: Association between forgiveness and revenge motivation



Exploratory analysis: Correlations robustness checks

The variables violated assumptions of normality, and we therefore added Spearman's rho to Table 2 which were very similar to Pearson's correlations and supported the robustness of the associations.

Exploratory analysis: Correlations comparisons

We conducted correlation comparisons with the “cocor” R package and found partial support for Hypothesis 2c that forgiveness is more strongly associated with behavioral motivations (i.e., conciliation, avoidance, and revenge) than empathy. We found support for empathy being positively correlated to conciliation motivation, $r(261) = 0.51$, 95% CI [0.41, 0.59], $p < .001$, avoidance motivation, $r(261) = -0.51$, 95% CI [-0.60, -0.42], $p < .001$, but did not find support for empathy as associated with revenge motivation, $r(261) = -0.11$, 95% CI [-0.23, 0.01], $p = .08$. All of the ten correlation comparison approaches indicated that forgiveness is more strongly associated with avoidance and revenge than empathy, yet none of the ten comparison approaches indicated that forgiveness is more strongly associated with conciliation than empathy. We provided full analyses and results for the comparisons in the “Additional analyses and results” section in the supplementary.

Comparing replication to original findings

We replicated all of the supported findings of the target article. We summarized the replication results and their interpretations based on LeBel et al. (2019) in Table 9.

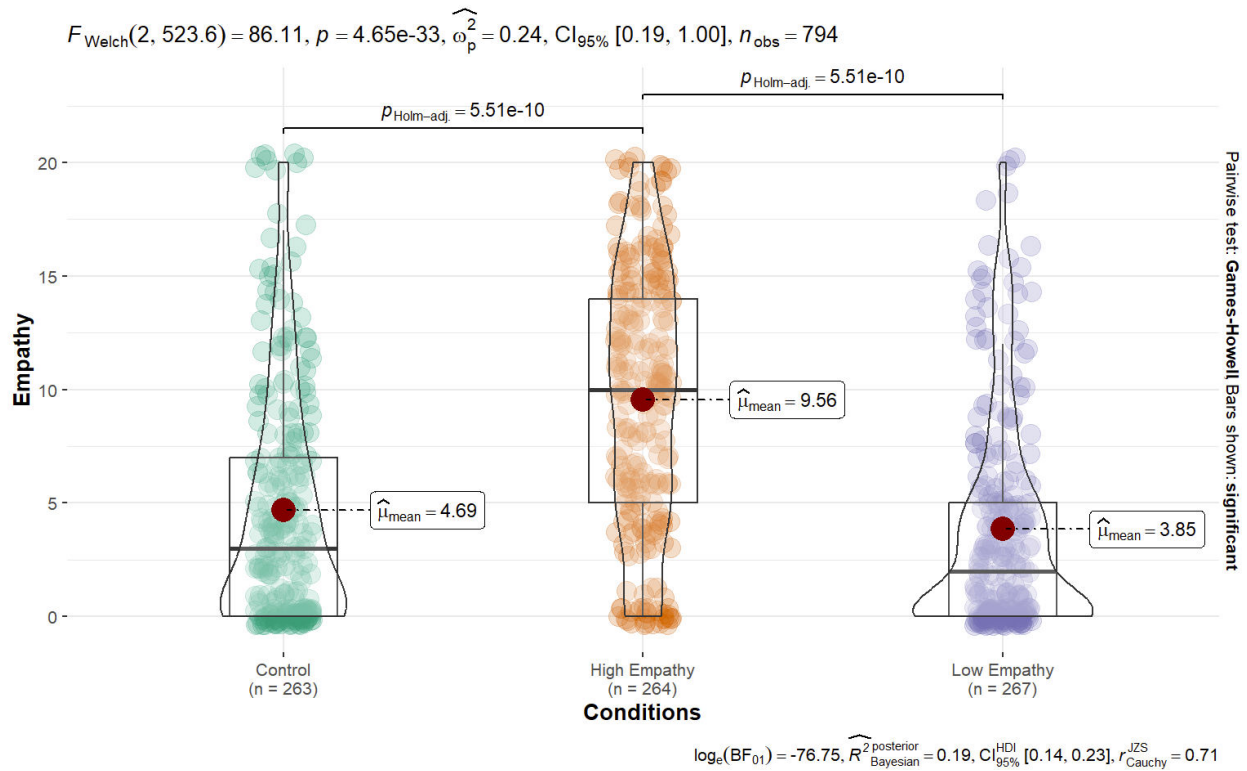
Extensions: empathy manipulation

Manipulation checks

We conducted independent samples Welch's t -tests (two-tailed) and found that affective empathy in the high empathy condition ($n = 264$; $M = 9.56$, $SD = 5.68$) was stronger than in the low empathy condition ($n = 267$; $M = 3.85$, $SD = 4.72$; $M_d = 4.84$; $t(509.74) = 12.58$, $p < .001$; $d = 1.09$, 95% CI [0.91, 1.27]), and the control condition ($n = 263$; $M = 4.69$, $SD = 5.24$; $M_d = 4.87$; $t(521.84) = 10.23$, $p < .001$; $d = 0.89$, 95% CI [0.71, 1.07]), indicating a successful manipulation (see Figure 7).

The differences in affective empathy between the low empathy condition and the control condition were weaker ($t(520.72) = -1.93$, $p = .050$; $d = 0.17$, 95% CI [-0.00, 0.34]), suggesting that for this recall task, when no instructions are given, people are more likely to recall low empathy situations.

Figure 7

Empathy (manipulation check): Comparison of empathy conditions

Note. Affective empathy scale is from 0 to 20, higher values indicate a stronger affective empathy towards the offender.

Forgiveness and apology

We conducted a one-way ANOVA and found support for empathy affecting forgiveness and differences in forgiveness across the three empathy conditions ($F(2, 791) = 32.8, p < .001; \eta_p^2 = 0.08, 90\% \text{ CI} [0.05, 0.11]$; See Figure 8).

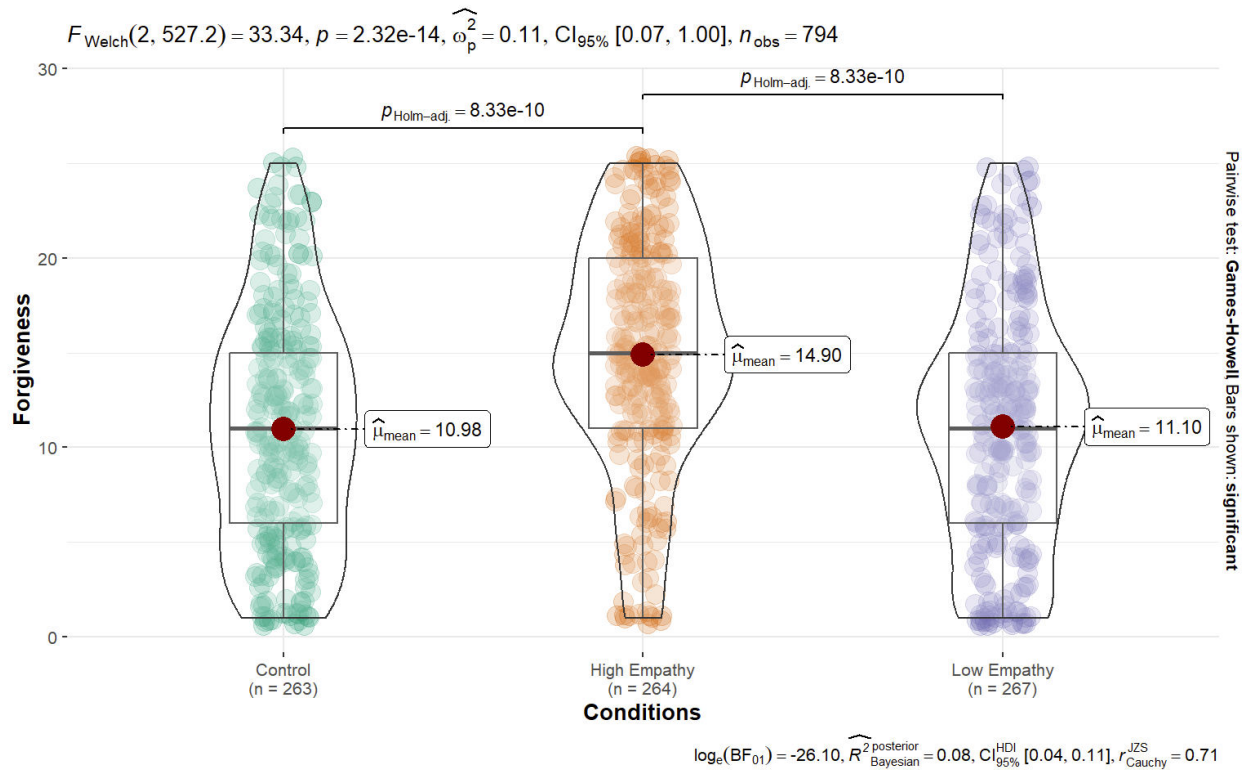
We conducted post-hoc Scheffe tests and found that forgiveness in the high empathy condition ($n = 264; M = 14.9, SD = 6.22$) was higher than in the low empathy condition ($n = 267; M = 11.1, SD = 6.4; M_d = 3.80, 95\% \text{ CI} [2.46, 5.15], p < .001; d = 0.60, 95\% \text{ CI} [0.43, 0.77]$), and

higher than in the control condition ($n = 263$; $M = 10.98$, $SD = 6.38$; $M_d = 3.92$, 95% CI [2.57, 5.28,], $p < .001$; $d = 0.62$, 95% CI [0.45, 0.79]). Yet, we did not find support for differences in forgiveness between the low empathy condition and the control condition ($M_d = 0.12$, 95% CI [-1.23, 1.47], $p = .97$; $d = -0.02$, 95% CI [-0.19, 0.15]).

We found support for differences in apology between the three empathy conditions ($F(2, 791) = 8.18$, $p < .001$; $\eta^2_p = .02$, 90% CI [0.01, 0.04]; See Figure 9).

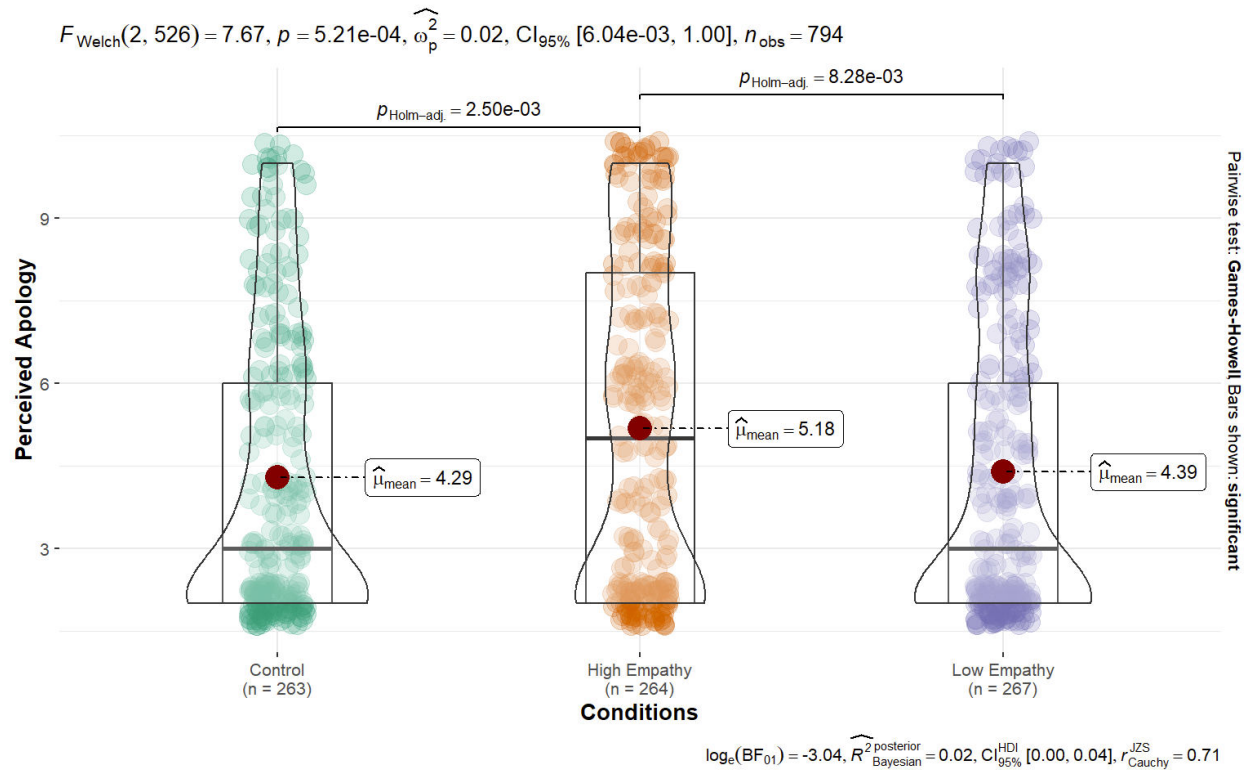
We conducted post-hoc Scheffe tests and found that perceived apology in the high empathy condition ($M = 5.18$, $SD = 2.94$) was higher than in the low empathy condition ($M = 4.39$, $SD = 2.7$; $M_d = -0.78$, 95% CI [-1.37, -0.20], $p = .005$; $d = 0.28$, 95% CI [0.11, 0.46]), and higher than in control condition ($M = 4.29$, $SD = 2.61$; $M_d = 0.89$, 95% CI [0.30, 1.47], $p = .001$; $d = 0.32$, 95% CI [0.15, 0.49]). We found no support for differences in perceived apology between the low empathy condition and the control condition ($M_d = 0.10$, 95% CI [-0.48, 0.68], $p = .916$; $d = -0.04$, 95% CI [-0.21, 0.13]).

Figure 8

Forgiveness: Comparison of empathy conditions

Note. Forgiveness scale is from 1 to 25, higher values indicate a stronger tendency to forgive the offender.

Figure 9

Perceived apology: Comparison of empathy conditions

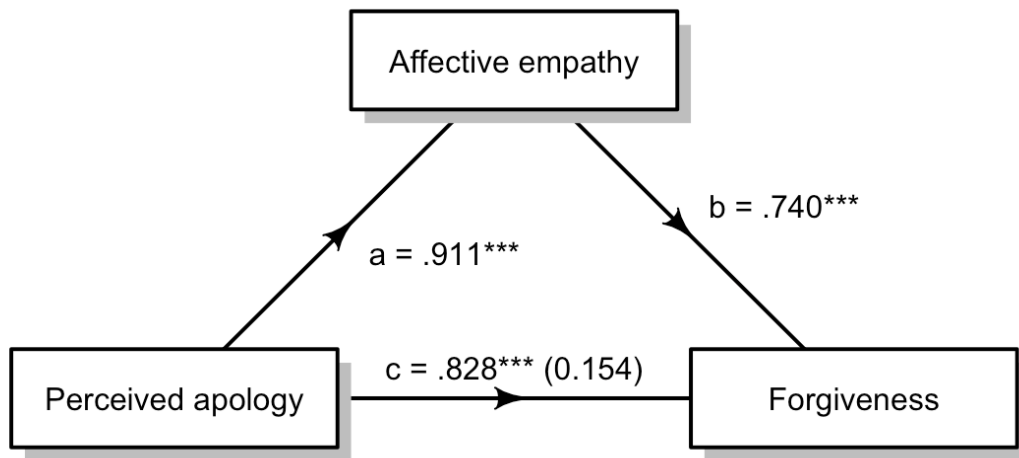
Note. Perceived apology scale is from 2 to 10, higher values indicate a stronger perceived apology from the offender.

Exploratory analysis: Mediation analyses

We conducted an exploratory mediation (bootstrapping) test to examine the meditation effects of empathy between apology and forgiveness in the control condition. We summarized the effects in Figure 10.

We found support for the effect of perceived apology on forgiveness as being mediated via affective empathy. Examining the regression coefficients, we found an association between perceived apology and forgiveness ($\beta = 0.83$, $t(261) = 5.82$, $p < .001$), between empathy and forgiveness ($\beta = 0.74$, $t(260) = 11.32$, $p < .001$), and between apology and empathy ($\beta = 0.91$, $t(261) = 8.24$, $p < .001$). The average direct effect was 0.15, 95% CI [-0.13 to 0.45], $p = .3.$, whereas the bootstrapped unstandardized indirect effect (Average Causal Mediation Effect, ACME) was 0.67, 95% CI [0.47 to 0.89], $p < .001$. Therefore, we found support for an indirect effect of affective empathy.

Figure 10

Exploratory mediation analyses in the control condition

Note. a b and c are regression coefficients between variables. *** $p < .001$.
(Average direct effect).

Discussion

We conducted a replication and extensions Registered Report of the empathy model of forgiveness. Our results are consistent with the findings reported in the target article (see Table 9 for a summary).

Replication

Overall, we found that: (1) affective empathy toward the offender is positively associated with forgiveness, (2) perceived apology is positively associated with empathy, (3) forgiveness is positively associated with conciliatory motivation and negatively associated with avoidance and revenge motivation, and (4) forgiveness is more strongly associated with behavioral motivation (i.e., conciliation, avoidance, and revenge) than is empathy.

These results are consistent with McCullough et al. (1997)'s conceptualization of forgiveness as empathy-facilitated motivational changes that promote relationship-constructive actions (i.e., conciliation) and inhibit relationship-destructive actions (i.e., avoidance and revenge) toward the offending person.

More than two decades after the original research was first published, the effect sizes are also remarkably similar and comparable to those in the target article, indicating the robustness and replicability of McCullough et al. (1997)'s model of forgiveness. Only two of the correlations from the hypotheses showed minor deviations from the target article's reported findings, with slightly stronger association between perceived apology and affective empathy (Original: $r = .36$; Replication: $r = .45$), and a weaker association between forgiveness and conciliatory motivation (Original: $r = .70$; Replication: $r = .45$).

Some often question the value of replication studies by stating that highly cited studies are self-evidently reliable and replicable. We believe these reactions often reflect a hindsight bias (a.k.a., knew-it-all-along phenomenon) that many, even the researchers, may hold towards replication studies. We previously demonstrated an ironic display of hindsight bias over the replicability of a classic experiment on hindsight bias (Study 3; Chen et al. 2021). To try and address hindsight bias over the replicability of our target article, we conducted a prediction poll on Twitter on March 26, 2023 (Feldman, 2023), and found that 21 out of 30 (70%) of the researchers in the community predicted an unsuccessful replication of McCullough et al. (1997)'s Study 1, lower than predictions for other targets included in the same Twitter poll. These predictions stand in strong contrast to the very successful replication we reported here, and further highlights the importance of testing intuitions and the possible misperceptions that some

may hold towards the importance of comprehensive independent Registered Reports of direct replication.

Extensions: Causality

We ran extensions examining the causal link of empathy on forgiveness and perceived apology. Our findings showed that: (1) affective empathy experienced by the wronged person contributes to interpersonal forgiveness, and (2) wronged persons who are empathetic towards the offender tend to perceive the offender as more apologetic. Overall, we found empirical support for our hypothesis that affective empathy is causally linked to forgiveness and perceived apology.

Although the causal link between empathy and forgiveness was proposed by the target article and is experimentally supported in our extension, the demonstrated impact of empathy on perceived apology is a new addition. Rather than a simple one-way cause-and-effect influence of perceived apology on affective empathy proposed in the original model, our results indicated that their relationship may be more complex. Affective empathy may also simultaneously influence the perceived apology in those situations. More research is needed to better understand the interplay between empathy and perceived apology and how it impacts the apology-forgiveness link.

As an exploratory analysis, we examined the mediating effect of empathy on the apology-forgiveness link. McCullough et al. (1997) hypothesized that affective empathy mediates the relationship between disposition on environmental variables and their causal effect on forgiveness. They tried to illustrate the mediational effect with structural equation modeling, whereas we explored its mediating effect with a bootstrapping test. Our findings in general supported the mediating effect of empathy on the apology-forgiveness relationship.

Implications, limitations, and directions for future research

We note several limitations. First, we followed the target article's methods and did not correct for multiple analyses and comparisons, which may potentially hamper the reliability of our results. We conducted many correlational tests and multiple ANOVAs with two dependent variables and pairwise comparisons, with some of the variables being not normally distributed, which may have heightened Type I error rate. Yet, we do not believe these have impacted our interpretation, as most of the effects were very large and the findings met a criteria of very low alpha ($p < .001$).

Second, we followed the target article's methods and did not employ any outlier handling strategies. As a close replication research study, we tried to ground our study on the original methodology as closely as possible to test replicability on the same grounds that the target article had to meet. Nevertheless, parametric tests, such as the Pearson correlations, may be sensitive to outliers (Knief & Forstmeier, 2021), and so we see room for discussion on whether replications should closely follow or aim to improve on the target's methods, even if not up to the best practices. We do not believe this had any impact on the results, as the highly similar effects of the target and the replication suggest.

Third, we note that our empathy manipulation in the extension was a manipulation of recalled past empathetic (vs. unempathetic) offending experience, but not the emotion that the participants experienced while participating in the experiment. This design is meant to align our extension with the methods used in the replication. There are important differences between manipulating the recalled past situation rather than the actual affect that participants experience. Future research may aim to supplement our methods and adopt more direct other empathy

manipulation techniques such as Batson et al. (1991)'s perspective-taking approach, contrasting objective versus emotional perspective towards the offenders.

We took the first step in manipulating empathy, and see promise for future research in also manipulating apology. Our extension revealed that affective empathy impacted perceived apology, yet it remains to be determined how perceived apology exactly affects and interacts with empathy, forgiveness, and other behavioral motivations. The link between perceived apology and forgiveness has been widely studied in the last decades (e.g., Hareli & Eisikovits, 2006; Schumann, 2012; Struthers et al., 2008), but few studies have investigated the link between perceived apology and empathy.

Following a successful replication of McCullough et al. (1997), we suggest future replications of important seminal follow-ups such as McCullough et al. (1998). McCullough et al. (1998)'s study, which was grounded and extended on our target article, has been an influential research paper contributing to our understanding of interpersonal forgiveness, and examining rumination, closeness, and revenge, which were incorporated into their original forgiveness model. At the time of writing (April 2023), there were 2603 Google Scholar citations of the research, with critical follow-up theoretical and empirical papers. We took initial steps to partially replicate some of the work by McCullough et al. (1998) by including the revenge measure in our replication, yet we see the value in a more comprehensive revisiting of their study. Importantly, as far as we know, there are no direct replications for the extended model of forgiveness.

Conclusion

Our replication of McCullough et al. (1997)'s Study 1 was successful. We found support for the empathy model of forgiveness, with affective empathy positively associated with perceived apology and forgiveness, forgiveness positively associated with conciliation and negatively associated with avoidance and revenge. In our extension, we demonstrated a direct causal link between empathy and forgiveness, and between empathy and perceived apology. Our exploratory extensions revealed an indirect mediational effect of affective empathy on the apology-forgiveness relationship.

References

- Agu, S. A., & Nwankwo, B. E. (2019). Influence of religious commitment, intentionality in marriage and forgiveness on marital satisfaction among married couples. *Ife Psychologia*, 27(2), 121–133.
- Akhtar, S., & Barlow, J. (2018). Forgiveness Therapy for the Promotion of Mental Well-Being: A Systematic Review and Meta-Analysis. *Trauma, Violence, & Abuse*, 19(1), 107–122. <https://doi.org/10.1177/1524838016637079>
- Batson, C. D., Charles D. (2011). Altruism in humans. Oxford University Press.
- Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic Joy and the Empathy-Altruism Hypothesis. *Journal of Personality and Social Psychology*, 61(3), 413–426. <https://doi.org/10.1037/0022-3514.61.3.413>
- Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imaging how you would feel. *Personality and social psychology bulletin*, 23(7), 751-758.
- Berenguer, J. (2007). The Effect of Empathy in Proenvironmental Attitudes and Behaviors. *Environment and Behavior*, 39(2), 269–283. <https://doi.org/10.1177/0013916506292937>
- Berenguer, J. (2010). The Effect of Empathy in Environmental Moral Reasoning. *Environment and Behavior*, 42(1), 110–134. <https://doi.org/10.1177/0013916508325892>
- Boon, S. D., Hojjat, M., Paulin, M., & Stackhouse, M. R. D. (2022). Between friends: Forgiveness, unforgiveness, and wrongdoing in same-sex friendships. *Journal of Social*

and Personal Relationships, 39(6), 1693–1716.

<https://doi.org/10.1177/02654075211062272>

Breitbart, W. (2018). Forgiveness. *Palliative & Supportive Care*, 16(3), 244–245.

<https://doi.org/10.1017/S1478951518000408>

Chen, E. Y., Chee, M. X., & Feldman, G. (2023). Revisiting the Differential Centrality of Experiential and Material Purchases to the Self: Replication and Extension of Carter and Gilovich (2012). *Collabra: Psychology*, 9(1). <https://doi.org/10.1525/collabra.57785>

Chen, J., Kwan, L. C., Ma, L. Y., Choi, H. Y., Lo, Y. C., Au, S. Y., ... & Feldman, G. (2021). Retrospective and prospective hindsight bias: Replications and extensions of Fischhoff (1975) and Slovic and Fischhoff (1977). *Journal of Experimental Social Psychology*, 96, 104154.

Chi, P., Tang, Y., Worthington, E. L., Chan, C. L., Lam, D. O., & Lin, X. (2019). Intrapersonal and interpersonal facilitators of forgiveness following spousal infidelity: A stress and coping perspective. *Journal of clinical psychology*, 75(10), 1896-1915.

Coles, N. A., Tiokhin, L., Scheel, A. M., Isager, P. M., & Lakens, D. (2018). The costs and benefits of replication studies. *The Behavioral and Brain Sciences*, 41, e124–e124. <https://doi.org/10.1017/S0140525X18000596>

Currier, J. M., Drescher, K. D., Holland, J. M., Lisman, R., & Foy, D. W. (2016). Spirituality, forgiveness, and quality of life: Testing a mediational model with military veterans with PTSD. *The International Journal for the Psychology of Religion*, 26(2), 167-179.

Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. PLoS ONE, 10(4): e0121945.

doi:10.1371/journal.pone.0121945

Donovan, L. A. N., & Priester, J. R. (2020). Exploring the Psychological Processes That Underlie Interpersonal Forgiveness: Replication and Extension of the Model of Motivated Interpersonal Forgiveness. *Frontiers in Psychology*, 11, 2107–2107.

<https://doi.org/10.3389/fpsyg.2020.02107>

Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, 64(325), 366-377.

Enright, R. D., & Coyle, C. T. (1998). Researching the process model of forgiveness within psychological interventions. *Dimensions of forgiveness: Psychological research and theological perspectives*, 1, 139-161.

Enright, R. D., & Fitzgibbons, R. P. (2000). Helping clients forgive: An empirical guide for resolving anger and restoring hope. American Psychological Association.

Enright, R. D., & Fitzgibbons, R. P. (2015). Forgiveness therapy an empirical guide for resolving anger and restoring hope (2nd ed.). American Psychological Association.

Fahimdanesh, F., Noferesti, A., & Tavakol, K. (2020). Self-Compassion and Forgiveness: Major Predictors of Marital Satisfaction in Young Couples. *The American Journal of Family Therapy*, 48(3), 221–234. <https://doi.org/10.1080/01926187.2019.1708832>

Farhadian, C. E., & Emmons, R. A. (2009). The psychology of forgiveness in the world religions. In *Forgiveness and reconciliation* (pp. 55-70). Springer, New York, NY.

Feldman, G. (2023). “Will it replicate? McCullough et al. 1997 | Study 1” poll on Twitter (Mar 26, 2023). Retrieved from <https://twitter.com/giladfeldman/status/1639889114630586370>

Feldman, G., Wong, K. F. E., & Baumeister, R. F. (2016). Bad is freer than good: Positive–negative asymmetry in attributions of free will. *Consciousness and Cognition*, 42, 26–40. <https://doi.org/10.1016/j.concog.2016.03.005>

Field, C., Zander, J., & Hall, G. (2013). Forgiveness is a present to yourself as well: An intrapersonal model of forgiveness in victims of violent crime. *International Review of Victimology*, 19(3), 235–247. <https://doi.org/10.1177/0269758013492752>

Field, S. M., Hoekstra, R., Bringmann, L., & van Ravenzwaaij, D. (2019). [When and Why to Replicate: As Easy as 1, 2, 3?](#). *Collabra: Psychology*, 5(1).

Fincham, F. D., Beach, S. R., & Davila, J. (2004). Forgiveness and conflict resolution in marriage. *Journal of family Psychology*, 18(1), 72.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4), 507-521.

Frise, N. R., & McMinn, M. R. (2010). Forgiveness and reconciliation: The differing perspectives of psychologists and Christian theologians. *Journal of Psychology and Theology*, 38(2), 83-90.

Gassin, E. A., & Lengel, G. J. (2014). Let me hear of your mercy in the mourning: forgiveness, grief, and continuing bonds. *Death studies*, 38(7), 465-475.

Giannini, H. C. (2017). Hope as Grounds for Forgiveness - A Christian Argument for Universal, Unconditional Forgiveness. *The Journal of Religious Ethics*, 45(1), 58–82.

<https://doi.org/10.1111/jore.12168>

Gordon, K. C., Hughes, F. M., Tomcik, N. D., Dixon, L. J., & Litzinger, S. C. (2009). Widening Spheres of Impact: The Role of Forgiveness in Marital and Family Functioning. *Journal of Family Psychology*, 23(1), 1–13. <https://doi.org/10.1037/a0014354>

Gottman, J. M. (1994). What predicts divorce? : the relationship between marital processes and marital outcomes. Lawrence Erlbaum Associates.

Guzmán, M., Tapia, M. J., Tejada, M., & Valenzuela, K. (2014). Evaluación del Perdón en Relaciones de Pareja: Propiedades Psicométricas del Transgression-Related Interpersonal Motivations Inventory (TRIM-18) en un Contexto Chileno. *Psyche*, 23(2), 1–12.

<https://doi.org/10.7764/psyche.23.2.623>

Hall, J., & Fincham, F. (2005). Self-forgiveness: The stepchild of forgiveness research. *Journal of Social and Clinical Psychology*, 24(5), 621–637.

<https://doi.org/10.1521/jscp.2005.24.5.621>

Hareli, S., & Eisikovits, Z. (2006). The role of communicating social emotions accompanying apologies in forgiveness. *Motivation and Emotion*, 30, 189-197.

Heirene, R. (2020, January 7). A call for replications of addiction research: Which studies should we replicate & what constitutes a “successful” replication?.

<https://doi.org/10.31234/osf.io/xzmn4>

Hendrickson, G. F., Stanley, J. C., & Hills, J. R. (1970). Olkin's new formula for significance of r_{13} vs. r_{23} compared with Hotelling's method. *American Educational Research Journal*, 7(2), 189-195.

Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of general psychology*, 130(2), 149-168.

Hotelling, H. (1940). The selection of variates for use in prediction with some comments on the general problem of nuisance parameters. *The Annals of Mathematical Statistics*, 11(3), 271-283.

Isager, P. M. (2019). [Quantifying Replication Value: A formula-based approach to study selection in replication research](#). *Open Science 2019, Trier, Germany*.

Jackson, J. C., Choi, V. K., & Gelfand, M. J. (2019). Revenge: A Multilevel Review and Synthesis. *Annual Review of Psychology*, 70(1), 319–345.
<https://doi.org/10.1146/annurev-psych-010418-103305>

Jeter, W. K., & Brannon, L. A. (2018). I'll Make It Up to You: Examining the effect of apologies on forgiveness. *The Journal of Positive Psychology*, 13(6), 597–604.
<https://doi.org/10.1080/17439760.2017.1291854>

Kelley, D. L., Waldron, V. R., & Kloeber, D. N. (2018). A communicative approach to conflict, forgiveness, and reconciliation: Reimagining our relationships. Routledge.

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Binan, E. D., Bocian, K., Busching, R., Cai, H., Cambier, F., Cantarero, K., Carmichael, C. L., Ceric, F., Chen, E. E., Cheong, W., ... Davis, W. E.

- (2018). Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science*, 1(4), 443–490.
<https://doi.org/10.1177/2515245918810225>
- Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, 53(6), 2576–2590.
- LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, 1, 389–402.
- LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3, 1–9. DOI: 10.15626/MP.2018.843
- Lee, Y.-R., & Enright, R. D. (2019). A meta-analysis of the association between forgiveness of others and physical health. *Psychology & Health*, 34(5), 626–643.
<https://doi.org/10.1080/08870446.2018.1554185>
- Macaskill, A. (2012). Differentiating dispositional self-forgiveness from other-forgiveness: Associations with mental health and life satisfaction. *Journal of Social and Clinical Psychology*, 31(1), 28.
- Maio, G. R., Thomas, G., Fincham, F. D., & Carnelley, K. B. (2008). Unraveling the Role of Forgiveness in Family Relationships. *Journal of Personality and Social Psychology*, 94(2), 307–319. <https://doi.org/10.1037/0022-3514.94.2.307>

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *The Behavioral and Brain Sciences*, 36(1), 1–15.

<https://doi.org/10.1017/S0140525X11002160>

McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Brown, S. W., & Hight, T. L. (1998). Interpersonal Forgiving in Close Relationships: II. Theoretical Elaboration and Measurement. *Journal of Personality and Social Psychology*, 75(6), 1586–1603.

<https://doi.org/10.1037/0022-3514.75.6.1586>

McCullough, M. E., Worthington, E. L., & Rachal, K. C. (1997). Interpersonal Forgiving in Close Relationships. *Journal of Personality and Social Psychology*, 73(2), 321–336.

<https://doi.org/10.1037/0022-3514.73.2.321>

McNulty, J. K. (2008). Forgiveness in Marriage: Putting the Benefits Into Context. *Journal of Family Psychology*, 22(1), 171–175. <https://doi.org/10.1037/0893-3200.22.1.171>

Mellor, D., Fung, S. W. T., & binti Mamat @ Muhammad, N. H. (2012). Forgiveness, Empathy and Gender—A Malaysian Perspective. *Sex Roles*, 67(1-2), 98–107.

<https://doi.org/10.1007/s11199-012-0144-4>

Menahem, S., & Love, M. (2013). Forgiveness in Psychotherapy: The Key to Healing: Forgiveness. *Journal of Clinical Psychology*, 69(8), 829–835.

<https://doi.org/10.1002/jclp.22018>

Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological bulletin*, 111(1), 172.

- Miller, A. J., Worthington, E. L., & McDaniel, M. A. (2008). Gender and Forgiveness: A Meta-Analytic Review and Research Agenda. *Journal of Social and Clinical Psychology*, 27(8), 843–876. <https://doi.org/10.1521/jscp.2008.27.8.843>
- Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. ., Keene, J. R., Protzko, J., Aczel, B., ... Baskin, E. (2018). The psychological science accelerator: advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, 1(4), 501–515. <https://doi.org/10.1177/2515245918797607>
- Nosek, B. A., & Errington, T. M. (2020). What is replication?. *PLOS Biology*, 18(3), e3000691. <https://doi.org/10.1371/journal.pbio.3000691>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., ber, A., Fidler, F., Hilgard, J., Struhl, M. K., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology*, 73(1), 719–748. <https://doi.org/10.1146/annurev-psych-020821-114157>
- Nouri, F. L., Lotfali, S., Sahranavard, S., Amiri, F., Fatideh, Z. A., & Fatideh, N. A. (2021). Measuring forgiveness among Iranian adolescents: Evaluation of psychometric properties of Persian version of transgression-related interpersonal motivations inventory. *Current Psychology*, 40(4), 1968–1978. <https://doi.org/10.1007/s12144-019-0135-5>
- Olkin, I. (1967). Correlations Revisited In: Stanley JC, editor. Improving Experimental Design and Statistical Analysis.

Olsson-Collentine, A., Wicherts, J. M., & van Assen, M. A. L. M. (2020). Heterogeneity in

Direct Replications in Psychology and Its Association With Effect Size. *Psychological Bulletin*, 146(10), 922–940. <https://doi.org/10.1037/bul0000294>

Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*, 6(61), 3167, doi:10.21105/joss.03167

Pearson, K., & Filon, L. N. G. (1897). Mathematical Contributions to the Theory of Evolution.

IV. On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation. *Proceedings of the Royal Society of London*, 62, 173-176.

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2022). That's a Lot to Process! Pitfalls of Popular Path Models. *Advances in Methods and Practices in Psychological Science*, 5(2), 251524592210958–. <https://doi.org/10.1177/25152459221095827>

Sandage, S. J., Long, B., Moen, R., Jankowski, P. J., Worthington, E. L., Wade, N. G., & Rye, M.

S. (2015). Forgiveness in the Treatment of Borderline Personality Disorder: A Quasi-Experimental Study: Forgiveness in the Treatment of Borderline Personality Disorder. *Journal of Clinical Psychology*, 71(7), 625–640. <https://doi.org/10.1002/jclp.22185>

Schumann, K. (2012). Does love mean never having to say you're sorry? Associations between relationship satisfaction, perceived apology sincerity, and forgiveness. *Journal of Social and Personal Relationships*, 29(7), 997-1010.

Sjöström, A., & Gollwitzer, M. (2015). Displaced revenge: Can revenge taste “sweet” if it aims at a different target? *Journal of Experimental Social Psychology*, 56, 191–202.

<https://doi.org/10.1016/j.jesp.2014.09.016>

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin*, 87(2), 245.

Strabbing, J. T. (2020). Forgiveness and Reconciliation. *Australasian Journal of Philosophy*, 98(3), 531-545.

Strelan, P., & Covic, T. (2006). A review of forgiveness process models and a coping framework to guide future research. *Journal of Social and Clinical Psychology*, 25(10), 1059-1085.

Struthers, C. W., Eaton, J., Santelli, A. G., Uchiyama, M., & Shirvani, N. (2008). The effects of attributions of intent and apology on forgiveness: When saying sorry may not help the story. *Journal of Experimental Social Psychology*, 44(4), 983-992.

Tarusarira, J. (2019). The Anatomy of Apology and Forgiveness: Towards Transformative Apology and Forgiveness. *The International Journal of Transitional Justice*, 13(2), 206–224. <https://doi.org/10.1093/ijtj/ijz006>

Thompson, L. Y., Snyder, C. R., Hoffman, L., Michael, S. T., Rasmussen, H. N., Billings, L. S., Heinze, L., Neufeld, J. E., Shorey, H. S., Roberts, J. C., & Roberts, D. E. (2005). Dispositional Forgiveness of Self, Others, and Situations. *Journal of Personality*, 73(2), 313–360. <https://doi.org/10.1111/j.1467-6494.2005.00311.x>

- Turnage, B. F., Hong, Y. J., Stevenson, A. P., & Edwards, B. (2012). Social Work Students' Perceptions of Themselves and Others: Self-Esteem, Empathy, and Forgiveness. *Journal of Social Service Research*, 38(1), 89–99. <https://doi.org/10.1080/01488376.2011.610201>
- Webb, J. R., & Toussaint, L. L. (2019). Forgiveness, well-being, and mental health. In *Handbook of forgiveness* (pp. 188-197). Routledge.
- Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 21(2), 396-399.
- Wong, L. C., Chu, A. M. Y., & Chan, C. L. W. (2014). Measuring the Transgression-Related Interpersonal Motivations Inventory in Marital Relationships: Forgiveness in the Chinese Context (C-TRIM). *Research on Social Work Practice*, 24(3), 356–363. <https://doi.org/10.1177/1049731513498622>
- Worthington, E. L., & Scherer, M. (2004). Forgiveness is an emotion-focused coping strategy that can reduce health risks and promote health resilience: Theory, review, and hypotheses. *Psychology & Health*, 19(3), 385-405.
- Worthington, E. L., Witvliet, C. V. O., Pietrini, P., & Miller, A. J. (2007). Forgiveness, health, and well-being: A review of evidence for emotional versus decisional forgiveness, dispositional forgivingness, and reduced unforgiveness. *Journal of behavioral medicine*, 30(4), 291-302.
- Xiao, Q., Yeung, S. K., Dunleavy, D. J., Röseler, L., Elsherif, M., & Feldman, G. (2023) Effect sizes and confidence intervals guide. DOI: 10.17605/OSF.IO/D8C4G . Retrieved from: <https://osf.io/d8c4g/>

- Yao, S., Chen, J., Yu, X., & Sang, J. (2017). Mediator Roles of Interpersonal Forgiveness and Self-Forgiveness between Self-Esteem and Subjective well-Being. *Current Psychology*, 36(3), 585–592. <https://doi.org/10.1007/s12144-016-9447-x>
- Yeung, S. K., & Feldman, G. (2022). Revisiting the Temporal Pattern of Regret in Action Versus Inaction: Replication of Gilovich and Medvec (1994) With Extensions Examining Responsibility. *Collabra: Psychology*, 8(1). <https://doi.org/10.1525/collabra.37122>
- Yu, L., Gambaro, M., Song, J. Y., Teslik, M., Song, M., Komoski, M. C., Wollner, B., & Enright, R. D. (2021). Forgiveness therapy in a maximum-security correctional institution: A randomized clinical trial. *Clinical Psychology and Psychotherapy*, 28(6), 1457–1471. <https://doi.org/10.1002/cpp.2583>
- Záhorcová, L., Enright, R., & Halama, P. (2021). The Effectiveness of a Forgiveness Intervention on Mental Health in Bereaved Parents—A Pilot Study. *Omega: Journal of Death and Dying*, 3022282110244–302228211024479. <https://doi.org/10.1177/00302228211024479>
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological methods*, 12(4), 399.

The link between Empathy and Forgiveness:
Replication and extensions of McCullough et al. (1997)'s Study 1

Supplementary

Contents

Open Science disclosures	2
Data collection	2
Conditions reporting	2
Variables reporting	2
Analysis of the original article	3
Effect size calculations of the original study effects	3
Power analysis of original study effect to assess required sample for replication	3
Sensitivity analysis	3
Materials and scales used in the replication + extension experiment	3
Exclusion criteria and handling outliers strategy	4
Comparisons and deviations	5
Pre-registration plan versus final report	5
Additional information about the study	7
Additional Tables and Figures	8
Definition of Empathy	8
Correlations comparisons	8
Replication evaluation	9
Replication closeness	9
Replication versus the original	10
References	11

Open Science disclosures

Data collection

Data collection was completed before analyzing the data.

Conditions reporting

All collected conditions are reported.

Variables reporting

All variables collected for this study are reported and included in the provided data.

Analysis of the [original article](#)

Effect size calculations of the original study effects

The effect size and confidence intervals for the intercorrelations between variables in the original study are summarized in Table 1, followed by the code blocks (Blocks 1-10) for R (Version 1.4.1717) computations and their output correspondingly (Figures 4- 13). Package of "MBESS" was adopted for the 95% confidence interval in the calculation in R.

See files Effect_size_and_power_analysis_for_original_study.Rmd/html in the OSF folder.

Power analysis of original study effect to assess required sample for replication

We summarized the power analyses for each correlational test in table 6 below. The required sample size to detect all the effects is 94 for power 0.95 and alpha 0.05, which is the power analysis for the correlation between apology and empathy. It has the highest number of minimum required N with significant effect.

See files Effect_size_and_power_analysis_for_original_study.Rmd/html in the OSF folder.

Sensitivity analysis

Three conditions between-subject one-way factor ANOVA, $N = 717$:

F tests - ANOVA: Fixed effects, omnibus, one-way

Analysis: Sensitivity: Compute required effect size

Input:	α err prob	=	0.05
	Power (1- β err prob)	=	0.95
	Total sample size	=	717
	Number of groups	=	3
Output:	Noncentrality parameter λ	=	15.5082455
	Critical F	=	3.0083367
	Numerator df	=	2
	Denominator df	=	714
	Effect size f	=	0.1470692

Post-hoc comparison between two groups, $n_1 = 239$, $n_2 = 239$:

t tests - Means: Difference between two independent means (two groups)

Analysis: Sensitivity: Compute required effect size

Input:	Tail(s)	=	Two
	α err prob	=	0.05
	Power (1- β err prob)	=	0.95
	Sample size group 1	=	239
	Sample size group 2	=	239
Output:	Noncentrality parameter δ	=	3.6121148
	Critical t	=	1.9649602
	Df	=	476
	Effect size d	=	0.3304286

Materials and scales used in the [replication + extension experiment](#)

See files RRR_McCullough_etal_1997_rep_ext.QSF/DOCX/PDF in the OSF folder.

Exclusion criteria and handling outliers strategy

We ran our analyses on the full sample of all those who completed the study successfully and answered all questions. Those who dropped out were not included.

We pre-registered that in the case of a failed replication, as a supplementary analysis and to examine any potential issues, we would also determine further findings reports with exclusions.

[We wrote: “In any case, we would report exclusions in detail with results for the full sample and results following exclusions (in either the manuscript or the supplementary. Criteria:

1. Participants indicating a low proficiency in English (self-report < 5, on a 1-7 scale)
2. Participants who self-report not being serious about filling in the survey (self-report < 4, on a 1-5 scale).
3. Participants who indicated issues or having seen these materials before in the funneling section (manually coded).”]

We also pre-registered that if we fail to find any support for our hypotheses, we would run statistical analyses after excluding the “outliers” with compensation for alpha (i.e. $\alpha = .01$) to account for multiple tests, but this part of our analysis would be purely exploratory since it would deviate from our initial study protocol.

Exclusions and outlier analyses were not carried out given our successful replication of the target article’s findings. We therefore followed the pre-registered plan to include all data in our analyses, without identifying or excluding outliers or applying additional exclusion criteria.

Comparisons and deviations

Pre-registration plan versus final report

See [Preregistration Planning and Deviation Documentation \(PPDD\)](#) document for the latest updates.

Components in your preregistration (e.g., stopping rule, analyses, hypotheses, exclusion rules)	Location of 1) preregistered decision/plan and 2) rational for decision/plan [Location / link]	Were there deviations? What type? [no / minor / major]*	If yes - describe details of deviation(s) [brief description / location / link]	Rationale for deviation [brief description / location / link]	How might the results be different if you had/had not deviated [brief description / location / link]	Date/time of decision for deviation + stage	Any additional notes
Study design	“Method” section –“Design and Procedure” of the main manuscript	no	/	/	/	/	/
Measured variables	“Method” section –“Measures” of main manuscript & Data_analysis.Rmd - “Compute variables”	minor	Inverted the score of 2 items in the forgiveness measure in R	The computation of forgiveness measure was mistaken.	The forgiveness measure was wrong.	9/4/2023: During data analysis.	Specific Items inverted: <i>I disapprove of him/her, I condemn the person.</i>
Demographics	“Method” section –“Participants” of main manuscript	minor	Deleted the row of age range.	Some participants indicated insensible age, e.g., 2, 7	Readers may get confused with the age range, i.e., 2-99	9/4/2023: During data analysis	/
Exclusion criteria	“Exclusion criteria” section of supplementary manuscript	no	/	/	/	/	/

Components in your preregistration (e.g., stopping rule, analyses, hypotheses, exclusion rules)	Location of 1) preregistered decision/plan and 2) rational for decision/plan [Location / link]	Were there deviations? What type? [no / minor / major]*	If yes - describe details of deviation(s) [brief description / location / link]	Rationale for deviation [brief description / location / link]	How might the results be different if you had/had not deviated [brief description / location / link]	Date/time of decision for deviation + stage	Any additional notes
IV	“Method” section -Table 4 of main manuscript	no	/	/	/	/	/
DV	“Method” section -Table 4 of main manuscript	no	/	/	/	/	/
Data analysis	Method” and “Results” sections of main manuscript	no	/	/	/	/	/

Note. *Categories for deviations: Minor - Change probably did not affect results or interpretations; Major - Change likely affected results or interpretation

Additional information about the study

1. Duration of Study Sessions: **7** days
2. No. of sessions: **4** (1st session is a pretest with 30 participants)
3. Time of Day: Participants were allowed to do the test at any time of their convenience.
4. Data collection dates: the data collection started on 27th March 2023 and ended on 3rd April 2023.
5. Participant Recruitment: Participants were recruited using Prolific.
6. No. of participants dropped out from the study: **136**

This study was conducted on Prolific with American participants. We imposed the following settings in recruiting our participants:

1. Participants were paid \$1.8 as a fixed participation reward. This amount was determined by multiplying the expected completion time (in mins.) with the minimal federal wage in the U.S. (i.e., \$0.125 per minute).
2. The expected completion time for each session was set at 12, 10, and 8 minutes in advance respectively.
3. The most time we allowed each worker to complete the study for each session was 49, 44, and 44 minutes respectively.
4. We limited all workers' HIT Approval Rate to be between 92% and 100%.
5. We limited each worker's number of HITs approved to be between 5,000 and 100,000.
6. We blocked Suspicious Geocode Locations and Universal Exclude List Workers.
7. We blocked duplicate IP addresses and duplicate geolocation.
8. We enabled HyperBatch so that all eligible workers were able to participate in our HIT immediately after the survey was launched.
9. We restricted workers' location to be in the U.S.
10. We did not restrict answering from specific devices

Additional Tables and Figures

Definition of Empathy

We used ChatGPT to summarize the meaning of “empathy” and “No empathy” from Michael McCulloughs’ works for our participants. The following is the full prompts and responses from ChatGPT.

Prompt:

"Please write a one paragraph description of a person showing "empathy" that in accordance with the academic literature on empathy based on the work of Michael McCullough."

ChatGPT:

"Empathy is the ability to understand and share the feelings of others. According to the work of Michael McCullough and other researchers in the field of empathy, an empathetic person may exhibit several characteristics. They may have a heightened ability to accurately recognize emotions in others and have an intuitive sense of what another person is feeling. They may also feel genuine concern for others and be willing to help them in a compassionate and non-judgmental way. Additionally, an empathetic person may be able to put themselves in another person's shoes and imagine what it might be like to experience their emotions, which can help them to better understand and connect with the other person. Overall, an empathetic person is likely to demonstrate a range of behaviors that reflect their ability to be sensitive to the emotional experiences of others and respond in a caring and supportive manner."

Prompt:

"Please write a one paragraph description of a person showing "NO empathy" that in accordance with the academic literature on empathy based on the work of Michael McCullough."

ChatGPT:

"According to the work of Michael McCullough and other researchers in the field of empathy, a person who lacks empathy may exhibit several characteristics. They may struggle to accurately recognize emotions in others and have difficulty understanding what another person is feeling. They may also lack concern for others and be unwilling to help them in a compassionate and non-judgmental way. Additionally, a person who lacks empathy may struggle to imagine what it might be like to experience the emotions of another person, which can make it difficult for them to understand or connect with others on an emotional level. Overall, a person who lacks empathy is likely to demonstrate a range of behaviors that reflect their inability to be sensitive to the emotional experiences of others and respond in a caring and supportive manner."

Prompt:

"Please help simplify the last paragraph to two sentences that would help any layperson understand the essence of lacking empathy."

ChatGPT:

"A person lacking empathy may struggle to understand and relate to the emotions of others, leading them to appear indifferent or unsympathetic to the needs and feelings of those around them. They may have difficulty recognizing and responding appropriately to the emotional cues and needs of others."

Correlations comparisons

Please see Data_analysis[version].Rmd/html for results of comparison of correlations.

Replication evaluation

Replication closeness

Lebel, McCarthy, Earp, Elson, and Vanpaemel (2018):

Target similarity		Highly similar				Highly dissimilar	
Category	Direct replication			Conceptual replication			
Design facet	Exact replication	Very close replication	Close replication	Far replication	Very far replication		
Effect/hypothesis	Same/similar	Same/similar	Same/similar	Same/similar	Same/similar		
IV construct	Same/similar	Same/similar	Same/similar	Same/similar	Different		
DV construct	Same/similar	Same/similar	Same/similar	Same/similar	Different		
IV operationalization	Same/similar	Same/similar	Same/similar	Different			
DV operationalization	Same/similar	Same/similar	Same/similar	Different			
Population (e.g. age)	Same/similar	Same/similar	Same/similar	Different			
IV stimuli	Same/similar	Same/similar	Different				
DV stimuli	Same/similar	Same/similar	Different				
Procedural details	Same/similar	Different					
Physical setting	Same/similar	Different					
Contextual variables	Different						

Figure X. Criteria for evaluation of replications by LeBel et al. (2018).

A classification of relative methodological similarity of a replication study to an original study. “Same” (“different”) indicates the design facet in question is the same (different) compared to an original study. IV = independent variable. DV = dependent variable.

“Everything controllable” indicates design facets over which a researcher has control.

Procedural details involve minor experimental particulars (e.g., task instruction wording, font, font size, etc.).

"Similar" category was added to the Lebel et al. (2018) typology to refer to minor deviations or extensions aimed to adjust the study to the target sample that are not expected to have major implications on replication success. See Olsson-Collentine, van Assen, and Wicherts (2020) on meta-analysis showing minor to no expected impact due to variations in sample population or setting.

Replication versus the original

LeBel, Vanpaemel, Cheung, and Campbell (2019) criteria:

A Signal Detected in Original Study

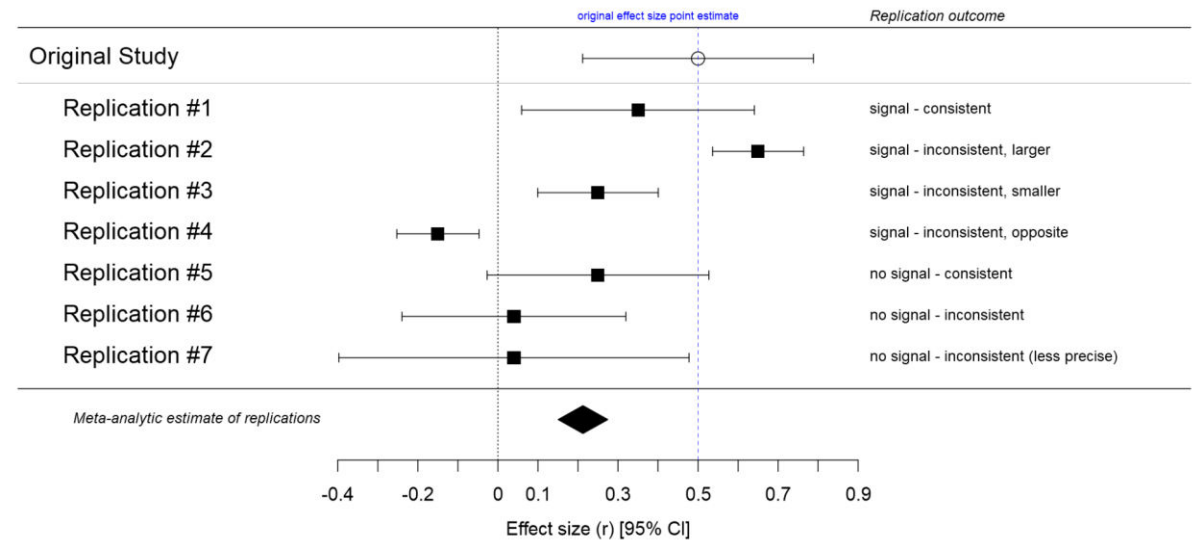


Figure X. Interpretation criteria for evaluation of replications outcomes by LeBel et al. (2019), if the original study detected a signal. A simplified replication taxonomy for comparing replication effects confidence intervals to target article original effect sizes.

B Signal Not Detected in Original Study

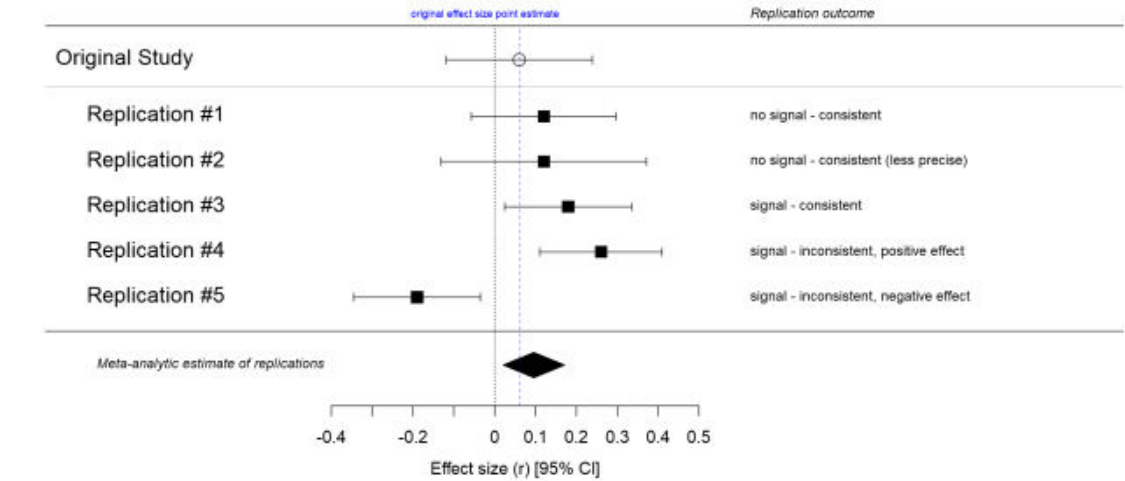


Figure X. Interpretation criteria for evaluation of replications outcomes by (LeBel et al., 2019), if the original study failed to detect a signal.

References

- Ben-Shachar, M. S., Lüdtke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, 5(56), 2815.
- Buchanan E, Gillenwaters A, Scofield J, Valentine K (2019). _MOTE: Measure of the Effect: Package to assist in effect size calculations and their confidence intervals_. R package version 1.0.2, <URL: <http://github.com/doomlab/MOTE>>.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & De Rosario, H. (2018). Basic functions for power analysis. *R Package Version* <https://cran.r-project.org/web/packages/pwr/pwr.pdf>.
- Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, 10(4): e0121945. doi:10.1371/journal.pone.0121945
- Hadley Wickham and Evan Miller (2021). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.4.3. <https://CRAN.R-project.org/package=haven>
- Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7. <https://CRAN.R-project.org/package=dplyr>
- John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Karline Soetaert (2020). diagram: Functions for Visualising Simple Graphs (Networks), Plotting Flow Diagrams. R package version 1.6.5. <https://CRAN.R-project.org/package=diagram>
- Ken Kelley (2021). MBESS: The MBESS R Package. R package version 4.8.1. <https://CRAN.R-project.org/package=MBESS>
- LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, 1, 389-402.
- LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, 3.
- Lüdtke D (2019). _esc: Effect Size Computation for Meta Analysis (Version 0.5.1)_.doi: 10.5281/zenodo.1249218 (URL: <https://doi.org/10.5281/zenodo.1249218>), <URL: <https://CRAN.R-project.org/package=esc>>.
- Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*, 6(61), 3167, doi:10.21105/joss.03167
- Signorell, A., Aho, K., Alfons, A., Anderegg, N., Aragon, T., Arppe, A., ... & Borchers, H. W. (2019). DescTools: Tools for descriptive statistics. *R package version 0.99*, 28, 17.