Forgiveness is generally described as the replacement of negative responses with positive ones towards a transgressor (see, for example, Enright & Fitzgibbons, 2000; Worthington, 2001; McCullough et al., 1998). What is necessary for forgiveness to occur? From a social-cognitive perspective, the literature emphasizes two points: (i) the offender must do something to encourage forgiveness, such as making amends, which encompasses sincere apology, remorse, atonement and taking responsibility for one’s actions (see, for example, Eaton, Struthers, & Santelli, 2006; McCullough et al., 1998; Zechmeister, Garcia, Romero, & Vas, 2004); and (ii) the victim needs to engage in some form of re-evaluation, such as downplaying the offender’s culpability and intentionality (Weiner, Graham, Peter, & Zmuidinas, 1991), recognizing the value of his or her relationship with the offender (McCullough, 2008), developing empathy for the offender (which facilitates forgiveness). Study 1 reveals that participants were more likely to forgive a friend’s negligence after being primed with punishment than after being primed with inability to punish. In Study 2, participants were more forgiving towards a criminal offender if the offender was punished by a judge than if the offender escaped punishment, a finding that was mediated by the just deserts motive. Study 3 was in the context of actual recalled ongoing interpersonal relations and revealed that punishment predicted forgiveness indirectly via just deserts, not via victims’ vengeful motivations. It is concluded that punishment facilitates forgiveness because of its capacity to restore a sense of justice. Copyright © 2013 John Wiley & Sons, Ltd.

Others have theorized about the potential relation between retribution and forgiveness (see Tripp, Bies, and Aquiono’s (2007) Vigilante Model of Justice; Worthington’s (2001) concept of the injustice gap; Fitness and Peterson’s (2008) perspective from evolutionary theory; and Exline, Worthington, Hill, and McCullough’s (2003) review of different social-cognitive predictors of forgiveness, such as apology, that could be construed as possessing justice-like characteristics). However, we are the first to attempt a systematic empirical (and predominantly experimental) investigation of the fundamental link between punishment and forgiveness. To our knowledge, only two previous studies have explicitly related punishment to forgiveness. David and Choi (2006) conducted a survey of Czechs who had been political prisoners under the Communist regime. They found that prisoners who knew that their tormentors had been condemned (i.e. punished) were more likely to forgive. Fitness and Peterson (2008) observed on the basis of interviews with married couples that punishment was ‘an integral part of the process of forgiveness’ (p. 262) in happy relationships.

Punishment, by definition, means that offenders themselves suffer in some way for what they have done (Carlsmith, Darley, & Robinson, 2002). Yet forgiveness is often considered a compassionate, loving act (e.g. Enright & Fitzgibbons, 2000; Worthington, 2001)—or, at least, a deliberate decision to not hold a grudge or not respond negatively to an offender (e.g. Yamhure Thompson et al., 2005). Indeed, a construct closely related to punishment, revenge, is often measured as the opposite of forgiveness (see the widely used Transgression Related Interpersonal Motivations [TRIM] scale of McCullough et al. (1998)). Thus, at first glance, it is somewhat paradoxical to speak of punishment as a facilitator of forgiveness. To

Research article

Retribution and forgiveness: The healing effects of punishing for just deserts

PETER STRELAN1* AND JAN-WILLEM VAN PROOIJEN2

1University of Adelaide, Adelaide, Australia; 2VU University, Amsterdam, The Netherlands

Abstract

Although punishment and forgiveness are considered to be opposites, in the present paper we propose that victims who punish their offender are subsequently more likely to forgive. Notably, punishment means that victims get justice (i.e. just deserts), which facilitates forgiveness. Study 1 reveals that participants were more likely to forgive a friend’s negligence after being primed with punishment than after being primed with inability to punish. In Study 2, participants were more forgiving towards a criminal offender if the offender was punished by a judge than if the offender escaped punishment, a finding that was mediated by the just deserts motive. Study 3 was in the context of actual recalled ongoing interpersonal relations and revealed that punishment predicted forgiveness indirectly via just deserts, not via victims’ vengeful motivations. It is concluded that punishment facilitates forgiveness because of its capacity to restore a sense of justice. Copyright © 2013 John Wiley & Sons, Ltd.

*Correspondence to: Peter Strelan, University of Adelaide, Adelaide, Australia.
E-mail: peter.strelan@adelaide.edu.au
substantiate our claim, we begin by elaborating on the different psychological functions that punishment serves.

Punishment is the behavioural representation of the construct of retributive justice (see, for example, Darley & Pittman, 2003). Retributive justice primarily serves a just deserts motive, which reflects a belief that punishment should occur for its own sake: Offenders must be seen to suffer, to a degree commensurate with their wrongful actions (Carlsmith, 2006; Carlsmith & Darley, 2008; Carlsmith et al., 2002; Darley, Carlsmith, & Robinson, 2000).

Transgressions stir emotions of anger, resentment, and condemnation and so prime a fundamental human need for balance and equity: That is, transgressors should not be allowed to get away with what they did (see, for example, theories of balance (Heider, 1958), deservingness (Feather, 1999) and equity (Adams, 1965)). As such, punishment may have the effect of returning relations between victim and offender to an even keel, insofar as a victim’s suffering has been cancelled out by an offender’s subsequent suffering. In this way, punishment for just deserts restores to victims the consensual values, identity (Okimoto & Wenzel, 2011), power (Bies & Tripp, 1996), status, autonomy (Shnabel & Nadler, 2008), control (Baumeister, Stillwell, & Heatherton, 1994) and self-esteem (Heider, 1958) that were lost or damaged by the transgression. Commensurately, there is evidence that empowered victims (Karremans & Smith, 2010) and those with high self-esteem (Strelan, 2007) possess the confidence to risk being vulnerable again and forgive.

There may also be an affective component associated with punishment: Seeing offenders suffer for their actions helps victims feel better (e.g., De Quervain et al., 2004; Stillwell, Baumeister, & Del Priore, 2008; although see Carlsmith, Wilson, & Gilbert, 2008, for a salient exception). Accordingly, there is evidence that individuals with an enhanced sense of well-being are better equipped to forgive (Bono, McCullough, & Root, 2008). Finally, punishment potentially allays some perceived costs of forgiving: the relinquishment of legitimate claims to justice and reparation; recidivism; perceived victim weakness; and loss of power (Baumeister, Exline, & Sommer, 1998; Lamb & Murphy, 2002).

In summary, punishment means that victims get justice (i.e. offenders are seen to get their just deserts). Regardless of the psychological outcomes that punishment delivers to victims, punishment in a broad sense may have the effect of helping to restore in victims a psychological sense of security that in turn facilitates forgiveness.

Finally, any discussion of justice and forgiveness should note that when justice is operationalized in inclusive as opposed to retributive terms, it is likely to be associated with forgiveness. For example, social justice cognitions (Karremans & Van Lange, 2005), just world beliefs for the self (Strelan & Sutton, 2011), rehabilitative punishment goals (Strelan, Feather, & McKee, 2011), procedural and distributive justice for the self (Lucas, Young, Zhdanova, & Alexander, 2010), and restorative justice (Strelan, Feather, & McKee, 2008; Wenzel & Okimoto, 2010) have all been positively linked to forgiveness.

However, in this article, our focus is on contrasting retribution—specifically, punishment—with situations where there either was no opportunity for punishment or the offender escaped punishment through other means. Retributive justice is important to people (Carlsmith et al., 2002) and normative in Westernized societies (Ellsworth & Gross, 1994), with even restorative justice processes requiring elements of retribution to be effective (Gromet & Darley, 2006). Indeed, people’s first port of call when a storm of injustice blows tends to be retributive, not inclusive. For example, physiological evidence suggests retaliation is an instinctual response to being transgressed against (Rusbult, Hannon, Stocker, & Finkel, 2005). Yet there are many occasions in people’s lives where they have been wronged and are unable to get retribution. We see it in intimate relationships where power imbalances prevent a subordinated partner from retaliating; we see it in the criminal justice system where individuals perceive that justice was not done on their behalf; and we see it occur for oppressed groups who have no access to justice.

In short, punishment plays a crucial role in regulating human behaviours and coexistence—yet the individual, interpersonal and group benefits of a contrary response, forgiveness, are also well established (see, for example, McCullough, 2008). Ironically, punishment has the potential to encourage forgiveness, but the effect may only become salient when we examine how people behave when they are unable to get retribution.

We tested relations between punishment and forgiveness across three studies, using three different methodological approaches. In Study 1, we primed punishment and tested its effect on forgiveness in the context of a friends’ negligence. In Study 2, we manipulated whether a criminal offender was punished by a judge in a hypothetical scenario and tested for the mediating effect of the just deserts motive. In Study 3, we utilized a correlational design in which participants recalled real-life experiences. There, our focus was on demonstrating the punishment–forgiveness link in ongoing interpersonal relationships and providing further evidence for the role of just deserts as an underlying process.

STUDY 1

Method

Study 1 employed a priming paradigm in which participants were randomly allocated to recall and describe in detail personally experienced transgressions where they either punished or were unable to punish their offenders. This type of approach has been employed to prime situational power (e.g., Anderson & Galinsky, 2006), and related techniques have been used in relation to forgiveness, priming for example justice (Karremans & Van Lange, 2005), power (Karremans & Smith, 2010) and relationships (Karremans & Aarts, 2007). One objective of the priming paradigm is to demonstrate that cognitive activation of one particular construct can affect another, seemingly unrelated construct (Galinsky, Gruenfeld, & Magee, 2003). In our case, we wanted to demonstrate that, fundamentally, people positively associate punishment with forgiveness. Following priming, participants indicated the extent to which they would forgive in an unrelated transgression scenario.
Participants

Originally, there were 52 participants, all undergraduate students at a large Australian city university who received course credit for participating. After screening (for details, see Results section), data from 49 participants were analysed (40 women, 9 men; M age = 22 years, SD = 7.75).

Procedures and Materials

Participants completed paper and pencil measures in individual cubicles. The study was advertised as being two separate pilot studies (‘Experiences of real life events’ and ‘Scenario piloting, Semester 2, 2011’) for different researchers from the same lab. To encourage the belief that the two studies were unrelated, different principal researcher names were attached to the advertised descriptions of each pilot study; different font was used for each set of study materials; and at the end of the priming section (‘Experiences of real life events’), participants read the following: ‘This is the end of the study. Thank you for taking the time to participate’. Debriefing at the end of the session revealed that no participant suspected the two studies were in fact related.

In the ostensible first study, instructions were as follows:

Please recall a situation in which you were deeply upset by another person’s actions and you punished that person [but you were unable to punish that person]. Please describe this situation—what happened and how you felt. It is important that you imagine this situation as vividly as possible.

After writing about the incident, participants completed items relating to the transgression itself: ‘How close were you to the person who upset you?’ (1 = not at all close, 7 = extremely close); ‘How upsetting was the event?’ (1 = not at all upsetting, 7 = extremely upsetting); and ‘Compared to all the other upsetting events you have experienced in your life, how upsetting was this one?’ (1 = least upsetting, 7 = most upsetting). We measured these variables because of the recall nature of the priming task and the fact that relational closeness and transgression severity are important predictors of forgiveness (Tsang, McCullough, & Fincham, 2006; Fincham, Jackson, & Beach, 2005); thus, we needed to control for them as potential confounds.

To cover the possibility that participants’ emotions in relation to their recalled transgression influenced responses, participants completed seven mood items (1 = not at all, 7 = completely in response to the tag, ‘Tell us how you feel right now about how you responded’): satisfied, angry, happy, resentful, content, annoyed and disappointed. After reverse coding, item scores were summed and averaged (α = .87).

Participants then went on to complete what was presented as the second pilot study. This in fact constituted the dependent measure, which was participants’ response to a hypothetical scenario adapted from Berry, Worthington, Parrot, O’Connor, and Wade’s (2001) Transgression Narrative Test of Forgiveness. Participants imagine they have recommended a friend to babysit a couple’s 3-year old child; the friend falls asleep, and the child drinks cleaning fluid and ends up in hospital for 2 days. This scenario has been used before in priming studies in relation to forgiveness (see, for example, Karremans & Smith, 2010; Karremans & Van Lange, 2005). Participants indicated the extent to which they would forgive their friend in that situation (1 = definitely not forgive, 5 = definitely forgive).

Results

Manipulation Checks

Two independent raters read each transgression description to check the extent to which participants followed priming instructions. Raters met to discuss and resolve disagreements. All participants in the punish condition (100%; n = 26) correctly followed instructions, describing situations in which they punished; and 88% (n = 23) of participants in the unable-to-punish condition correctly followed instructions. Three participants from the latter condition were subsequently excluded from analyses because they described situations in which they in fact punished or it was not clear that they had been unable to punish.1

Main Analysis

An independent samples t-test revealed that, as predicted, participants in the punish condition (M = 3.08, SD = 0.93) were more likely to forgive than participants in the unable-to-punish condition (M = 2.35, SD = 0.88), t(47) = 2.79, p = .008, η² = 0.14.

Checking for Potential Confounds

Independent samples t-tests were also conducted for emotions, relational closeness, perceived transgression severity and comparative severity. There were no differences between conditions on emotions, t(47) = 1.07, p > .1, and perceived transgression severity, t(47) = 1.24, p > .1. However, participants in the punish condition were more likely to be close to their transgressor than those in the unable-to-punish condition, t(47) = 2.70, p = .010, and were less likely to rate the transgression as comparatively severe, t(47) = 2.79, p = .007. This suggested the possibility that relational closeness and/or the comparative severity of the transgression rather than the punishment priming could be explaining differences in forgiving. Consequently, we re-ran the main analysis as an analysis of covariance controlling for closeness and comparative severity. The outcome was the same, F(1, 45) = 6.05, p = .018, η² = 0.12.2

Thus, Study 1 provides initial support for our hypothesis that punishment relative to an inability to punish enables forgiveness. This finding is especially encouraging given that the priming of punishment predicted forgiveness in an unrelated context.

STUDY 2

We had three aims in Study 2. The first was to replicate our findings with a different methodological approach, specifically,

1Including these three participants in analyses does not alter results.
2Alternatively, it was possible that closeness and/or severity moderated the effect of condition. To test this, we mean-centred each of the closeness and severity variables, computed interaction terms with condition and conducted separate regressions for each. Neither interaction term contributed significant variance on forgiveness over and above the main effects of condition and closeness or severity (for closeness, p = .074; for severity, p = .90).
a hypothetical scenario embedded within an experimental design. Of course, hypothetical scenarios by themselves are limited: what people forecast they will do, think or feel is not always the same as what they actually do, think or feel (see, for example, Carlsmith et al., 2008). However, as part of a package of studies utilizing different methodological approaches, and specifically in the context of punishment and forgiveness, hypothetical scenarios can be valuable. Specifically, they enable one to set levels of transgression meaningfulness and punishment options in a way that would be ethically impossible in the laboratory and at the same time afford an ecological validity that is often absent within the necessarily artificial constraints of a lab-based transgression context.

The second aim was to examine the extent to which the punishment effect extends from the interpersonal to a third-party context. Thus, we positioned participants’ transgressive experiences within the criminal justice system, varying the extent to which an offender was punished or not by the courts. To the extent that punishment that is delivered by a legitimate judge satisfies a victim’s desire for justice, such third-party punishment may still facilitate forgiveness. Hence, we expected that even when punishment is indirectly delivered—that is, by the state on behalf of the victim—punishment will still encourage a forgiving response. This approach, in turn, allowed us to examine a nuance of situations where victims are unable to punish. Victims may have an opportunity to punish but still not be able to take it and therefore fail to punish; alternatively, victims may not have an opportunity to punish. Either way, the result is that the offender is not punished. Thus, in addition to a punishment condition, we had two no punishment conditions: ‘fail to punish’ and ‘unable to punish’.

The third aim of the study was to begin to examine the potential mediating effect of the just deserts motive on the relation between punishment and forgiveness. We have already noted that just deserts is the primary motivation behind punishment (e.g. Carlsmith & Darley, 2008). Accordingly, one of the main reasons why our line of reasoning predicts a positive effect of punishment on forgiveness is because of the just deserts motive, that is, the feeling that punishment reinstalls justice.

**Method**

**Participants and Design**

We recruited a total of 42 participants (13 men, 29 women; mean age = 21 years, SD = 2.49) by means of flyers in student cafeterias of a large university. They were assigned randomly to one of three punishment conditions (punishment vs. fail to punish vs. unable to punish). The study was followed by other, unrelated experiments. The battery of studies lasted approximately 20 minutes, and participants were paid €2.50.

**Procedure**

Upon entry in the laboratory, participants were seated in separate cubicles. We used computer equipment to present the stimulus materials and register the data. Participants were asked to imagine themselves in the following scenario:

At the end of a day’s work you are in the tram on your way home. It is crowded in the tram, so you are standing close to other people in the aisle. Suddenly you notice that a man tries to steal your wallet from your pocket. Before you are able to stop the thief, he has your wallet in his hand and hurries towards the exit of the tram. The thief manages to get out and run away.

In the punishment condition, participants then read the following:

Your report this incident to the police, after which the thief is caught. The wallet is never retrieved. The man has to appear in court. He is convicted and receive the maximum legal punishment for this type of criminal offense.

In the fail-to-punish condition, participants instead read the following information:

Your report this incident to the police, after which the thief is caught. The wallet is never retrieved. The man had to appear in court. Due to a lack of evidence he is acquitted and receives no punishment for his crime.

In the unable-to-punish condition, participants read the following information:

Your report this incident to the police, but the thief is never caught. The wallet is never retrieved.

We then assessed the dependent variables (DVS). To measure perceived forgiveness, participants indicated to what extent they agreed to the following three statements (1 = strongly disagree, 7 = strongly agree): ‘I would forgive the man for he did’, ‘I would want the man to suffer more because of this crime’ (recoded) and ‘If I would run into this man again I probably would no longer be upset’. These three items were averaged into a reliable forgiveness scale (α = .71). To measure participants’ perception of whether the offender received his just deserts, we assessed participants’ agreement to the following items (1 = strongly disagree, 7 = strongly agree): ‘The man got what he deserved’, ‘I believe that justice has prevailed in this incident’ and ‘This incident ended in a just way’. These items were averaged into a reliable justice scale (α = .90). Finally, we checked the manipulation by means of the following items (1 = strongly disagree, 7 = strongly agree): ‘The man was punished for the theft’, ‘The man experienced negative consequences of the theft’ and ‘The man had to suffer as a consequence of his acts’. These items were averaged into a reliable manipulation check scale (α = .90). Participants were then informed that the study had ended. They were debriefed, thanked and paid for their participation.

**Results**

**Manipulation Check**

We analysed the results with one-way analyses of variance (ANOVAs). The effect of our punishment manipulation on the manipulation check was significant, F(2, 39) = 19.05, p < .001. We then performed a Helmert contrast comparing the mean of the punishment condition (M = 5.57, SD = 1.49) with the means of the fail-to-punish (M = 2.33, SD = 1.71) and unable-to-
punish conditions ($M = 2.17$, $SD = 1.73$). This contrast was significant, $F(1, 39) = 38.03$, $p < .001$. These results indicate that participants believed that the thief was punished more strongly in the punishment condition than in the two conditions where no punishment took place, as intended with this manipulation. The contrast comparing the fail-to-punish with unable-to-punish condition was not significant, $F < 1$.

**Forgiveness**

The punishment manipulation exerted a significant influence on forgiveness, $F(2, 39) = 4.00$, $p < .03$. To test our hypothesis, we conducted a Helmert contrast. The contrast testing the punishment condition ($M = 3.67$, $SD = 1.00$) against the fail-to-punish ($M = 2.48$, $SD = 1.34$) and unable-to-punish conditions ($M = 2.60$, $SD = 1.31$) was significant, $F(1, 39) = 7.94$, $p < .01$. Consistent with our hypothesis, these results indicate that forgiveness is higher if the offender is punished than if the offender is not punished. The contrast testing the fail-to-punish condition against the unable-to-punish condition was not significant, $F < 1$.

**Perceived Justice**

The manipulation also had a significant effect on perceived justice, $F(2, 39) = 34.87$, $p < .001$. A Helmert contrast revealed that participants perceived more justice in the punishment condition ($M = 5.33$, $SD = 1.25$) than in the fail-to-punish ($M = 1.52$, $SD = 0.65$) and the unable-to-punish conditions ($M = 2.43$, $SD = 1.67$), $F(1, 39) = 66.13$, $p < .001$. The contrast comparing the fail-to-punish with the unable-to-punish condition was not significant, although there was a trend towards more positive justice perceptions in the unable-to-punish condition than in the fail-to-punish condition, $F(1, 39) = 3.60$, $p < .07$.

**Mediation Analysis**

To test whether perceived justice mediated the effects of the punishment manipulation on forgiveness, we inserted the vectors of the two orthogonal Helmert contrasts ($2 - 1 - 1$; $0 1 - 1$) in a series of regression analyses. In separate regression analyses, the first contrast had a significant effect on forgiveness ($B = 0.38$, $p < 0.01$) and on perceived justice ($B = 1.12$, $p < .001$), mirroring the previously reported ANOVA results. We then included perceived justice as independent variable (IV) in a regression analysis on forgiveness. Results revealed that the effect of the first contrast on forgiveness was reduced to nonsignificance ($B = 0.01$, $p = .97$), and the effect of perceived justice was significant ($B = 0.33$, $p < .04$). A bootstrapping analysis (5000 samples, bias-corrected) revealed that the indirect effect of punishment via justice was significant, $B = 0.37$, $CI_{95\%} = [0.05, 0.98]$. As to the second contrast, none of the conditions for mediation were met, so we did not proceed further.

In sum, punishment had an indirect effect on forgiveness via justice, thus meeting one of the primary aims of Study 2. These results suggest that punishment addresses the just deserts motive by promoting a sense of justice, which in turn facilitates forgiveness.

Another aim of the study was to generalize the Study 1 findings from first-person to third-party punishment. We should acknowledge, however, that the impetus for punishment in the context of the criminal justice system usually still lies with an individual. That is, justice may be formally delivered by a third-party on the victim’s behalf—but only because the victim reported a crime to the police and pressed charges in the first place. In other words, in the criminal justice context such as the one used here, an individual victim implicitly seeks punishment and therefore, technically, the third-party context is not ‘purely’ third party.

Finally, we note that demand characteristics may have been operating in this study, insofar as university undergraduates—who presumably are not naive about structural explanations for crime—may have been led to forgive. We dealt indirectly with this limitation, along with some others, in the third and final study.

**STUDY 3**

Having found a likely causal relationship between punishment and forgiveness, the main aim of the third study was to extend our results to the context in which forgiveness is most relevant: ongoing interpersonal relationships. We also wanted to demonstrate that the punishment–forgiveness relation occurs irrespective of methodology. Thus, Study 3 utilized a survey-type, recall design, standard in forgiveness research (e.g. McCullough, Hoyt, & Rachal, 2000). In our case, we asked all participants to recall and describe a situation in which they had been hurt and they subsequently took steps to punish the person who hurt them. Whereas the previous two studies examined the effects of punishment in comparison with conditions where the offender was not punished, in Study 3, all participants had punished.

Finally, we addressed a potential limitation: We have not yet accounted—either conceptually or methodologically—for the distinction between just deserts and revenge as independent retributive motives for punishment. Our focus in these studies is on just deserts as the primary motivating force for punishing—yet punishment is often also motivated by the desire for revenge (see, for example, Ho, ForsterLee, ForsterLee, & Crofts, 2002). Given that both just deserts and revenge involve victims wanting to see offenders suffer, how can we be sure that the participants in our studies are not conflating the two constructs?

Theoretically, at least, the line between retribution as just deserts and retribution as revenge is reasonably clear: In the case of the former, justice is restored through a proportionate response; in the case of the latter, victims not only want to re-balance the scales of justice but they also want to retaliate (Gerber & Jackson, 2013). Although both just deserts and revenge are shaped by negative emotions, punishment motivated by revenge is likely to possess a relatively heightened emotional intensity that in reality often renders a victim’s response over-compensatory (Gollwitzer, Meder, & Schmitt, 2011), often with the effect of encouraging counter-retaliation (Stillwell et al., 2008).
In short, both just deserts and revenge motives entail victims wanting to see their offenders suffer—but in the case of just deserts, victims perceive that the suffering should be equivalent to their own and has the ultimate purpose of restoring a sense of justice, whereas in the case of revenge, victims perceive that suffering should occur in addition to justice being done or even if justice is not done (for a detailed discussion of the distinction between just deserts and revenge goals of punishment, see Gerber & Jackson, 2013). Thus, the just deserts motive for punishment represents a fair response, whereas revenge does not. Accordingly, in Study 3, we measured both just deserts and revenge motives for punishing. We aimed to demonstrate, first, that punishment is positively associated with both just deserts and revenge motives. In turn, we expected that, to the extent that punishment is motivated by revenge, victims would be less likely to forgive (McCullough et al., 1998), whereas—consistent with our theorizing—to the extent that punishment is motivated by just deserts, victims would be more likely to forgive.

Method

Participants

Participants were members of the community (n = 46) and students from a large university participating for course credit (n = 64). Following data cleaning (see details later), data from 94 participants were analysed. Of these, there were 79 women and 15 men (M_{\text{age}} = 27 years, SD = 11.24; range 17–64).

Procedures and Materials

Community participants heard about the study after a research assistant advertised the study on various social media outlets (email; Facebook; online forums). Both community and student participants were directed to an online link to the study. They responded to the following items:

Participants were asked to

Please recall a situation from the last 12 months in which another adult did something that hurt or upset you, and you in turn did something to punish them for what they did. Punishment can be anything you did that was designed to get justice for yourself. This can take lots of forms and can be active or passive. For example, silent treatment; gossiping or spreading rumours; telling someone off so they feel bad; refusing to return their phone calls; de-friending them on Facebook, etc. Everyone responds differently. We are interested to know what you did in response to being hurt. The person who hurt you could have been anyone with whom you are still involved—a relationship partner, a family member, a friend, a work colleague, etc. Please try to recall the situation as vividly as possible. There are no right or wrong answers. Please be open and honest in your responses—the survey is totally anonymous. When you write about this situation, we will ask you to describe what the person did to you and how it made you feel; describe what you did to punish them; explain what you expected to achieve by punishing them; and explain why you think your response could be considered a punishment. Please continue now, by telling us what happened...

Punishment and Background Variables. We measured the extent to which participants believed they had punished their transgressor with the item, ‘I punished the other person for what he/she did’ (1 = completely disagree, 7 = completely agree); current relationship involvement with the item, ‘I cut off the relationship with him/her’ (1 = strongly disagree, 5 = strongly agree, subsequently recoded); and the severity of the initial transgression experienced by the victim with the mean of two items, ‘How hurtful were their actions?’ (1 = not at all upsetting, 7 = extremely upsetting) and ‘Compared with all the other hurtful events you have experienced in your life, how upsetting was this one?’ (1 = least upsetting, 7 = most upsetting; r = .73, p < .001). Participants also indicated time elapsed since the transgression, with responses subsequently coded into days.

Just Deserts Motive (1 = Completely Disagree, 7 = Completely Agree). Just deserts motive was measured with the mean of three items: ‘By punishing them, I made sure that...they did not get away with what they did; they got their just deserts; they hurt as much as I did’ (x = .75).

Revenge Motive. Revenge motive was measured with the five-item revenge subscale of the TRIM (McCullough et al., 1998; 1 = strongly disagree, 5 = strongly agree). Items were summed and averaged, with higher scores indicating desire for revenge (x = .89).

Forgiveness. Forgiveness was measured with a single item (‘I forgive him/her for what he/she did to me’, where 1 = strongly disagree, 5 = strongly agree) and also the five-item benevolence subscale of the TRIM (McCullough, Fincham, & Tsang, 2003) scale (1 = strongly disagree, 5 = strongly agree). Benevolence items were summed and averaged, with higher scores indicating greater benevolence (x = .87).

Results

Data Screening

Because the study was conducted online, we asked a number of questions to establish the validity of the responses provided (true/false format), specifically the following: ‘My answers are not valid’ (n = 3); ‘I completed the questionnaire with someone else’ (n = 0); ‘I was talking to other people during the study’ (n = 11); and ‘I was under the influence of alcohol or other substances’ (n = 0). Anyone who indicated ‘yes’ on at least one of these items was subsequently removed from analyses. A further two participants failed to complete a substantial portion of the survey and were also not included.

Punishment and Background Variables

In general, participants strongly agreed that they had punished the person who initially hurt them (M = 4.99, SD = 1.46). Mean time since the events around the initial transgression and subsequent punishment was approximately 6 months (M = 177 days, SD = 221 days), ranging from 1 day to 3 years (12 participants recalled transgressions that occurred beyond the specified ‘last 12 months’ period; we retained them for analysis because, as will be seen shortly, time elapsed since the transgression was not significantly associated with our key variables, and we
wanted to maximize power). Current level of relationship involvement was high ($M=4.38$, $SD=0.72$).

**Relations between Punishment and Forgiveness**

Relations between punishment, forgiveness-related and background variables are shown in Table 1. It may be seen that degree of punishment was positively associated with just deserts, approached significance in a positive direction with revenge, but was unrelated to forgiveness and benevolence. Of course, we had kept punishment essentially constant, and so the subsequent lack of variability on this measure may explain the null relations with forgiveness and benevolence. Alternatively, these data suggest that it is not the degree of punishment that matters but, rather, that punishment was administered at all. If punishment is administered, it is the motivation for punishing that becomes relevant to forgiveness. Accordingly, just deserts was positively associated with forgiveness and approached significance with benevolence and was not associated with revenge. Revenge was negatively associated with forgiveness and benevolence.

Finally, degree of original hurt was positively associated with revenge, but not associated with the just deserts motive and either of the forgiveness measures. Time elapsed since the transgression was not related to any of the variables. Relationship involvement was positively associated with forgiveness and benevolence, negatively associated with revenge and unrelated to just deserts.

We subsequently tested for the indirect effects of punishment on forgiveness via the just deserts and revenge motives. We employed Preacher and Hayes’ (2004) bootstrapping procedure (5000 iterations, bias-corrected). We entered punishment as the IV, and just deserts and revenge motives as mediators. Because forgiveness and benevolence were highly correlated ($r=.71$), we computed a composite score for these two variables and entered it as the DV (hereafter referred to as ‘forgiveness’). In this multiple mediation model, punishment did not exert a direct effect on forgiveness ($B=0.10$, $p=.44$). However, punishment was positively associated with the just deserts motive ($B=0.42$, $p<.001$) and the revenge motive ($B=0.10$, $p=.07$). In turn, just deserts was positively associated with forgiveness ($B=0.27$, $p=.014$), whereas revenge was negatively related ($B=-1.41$, $p<.001$). The indirect path of punishment influencing forgiveness through just deserts was significant, $B=0.11$, $CI_{95\%}=[0.03, 0.25]$. However, the indirect path of punishment influencing forgiveness through revenge was not significant, $B=-0.15$, $CI_{95\%}=[-0.31, 0.01]$.

These findings reveal that punishment exerts its positive effect on forgiveness only through the just desert motive. Although the indirect effect of punishment on forgiveness via revenge was not significant, revenge motivation had a strong negative effect on forgiveness. This raises the possibility that revenge motivation suppresses the direct effect of punishment on forgiveness. To explore this possibility, we conducted two additional bootstrapping analyses, one in which we controlled for revenge and one in which we controlled for just deserts. When revenge is controlled for, the relation between punishment and forgiveness becomes significant ($B=0.25$, $p=.02$), but this was not the case when just deserts was controlled for ($B=0.02$, $p=.90$). Thus, it appears that the revenge motive suppresses the effect of punishment on forgiveness.

In summary, Study 3 consisted of individuals who were maintaining an ongoing relationship with a person who had hurt them in the past. It appears that punishment has a positive indirect effect on forgiveness via the just deserts motive. Victims may be able to forgive after punishing, but only to the extent that punishment is motivated by a desire to get justice for themselves. The link between punishment and forgiveness is disrupted if punishment is motivated by a desire for vengeance. Finally, there is some evidence that punishment may be directly associated with forgiveness, but only after one accounts for victims’ vengeful motivations.

**GENERAL DISCUSSION**

These three studies provide evidence that punishment encourages forgiveness. Importantly, the link between punishment and forgiveness was found across three different methodologies (priming, hypothetical third party and personal recall) that featured two standardized transgressions (one by a friend and one by a stranger) and two different types of transgressors (a stranger and someone with whom the victim was in an ongoing relationship), and with three different dependent measures of forgiveness. Study 1 showed that when people are primed to recall a time when they punished a transgressor—as opposed to being unable to punish—they were more likely to forgive a friend in an unrelated context. Thus, at an abstract level, it appears that people positively associate the concepts of punishment and forgiveness. Study 2 demonstrated that the effect could be replicated in the impersonal context of the criminal justice system. When a third party (i.e. a judge) punished a stranger who transgressed, participants in the role of victims were more likely to forgive the stranger than if the stranger was not punished or there was no opportunity for punishment. In Study 3, the investigation moved from the abstract and the hypothetical third-party contexts to the explicitly personal. For victims in an ongoing relationship, punishment indirectly resulted in increased forgiveness when victims were motivated by just deserts but not revenge.

Having established that punishment has a causal effect on forgiveness, a secondary aim of this research was to test an explanatory mechanism. We postulated that punishment delivers justice to victims, and it is the getting of justice that enables forgiveness. If offenders are perceived to experience suffering equivalent to that of victims, then relations between victims
Table 1. Pearson product moment correlations between punishment and forgiveness-related and background variables (Study 3)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Punishment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Just deserts</td>
<td>.40**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Revenge</td>
<td>.19*</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Forgiveness</td>
<td>.07</td>
<td>.24**</td>
<td>.53**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Benevolence</td>
<td>.19*</td>
<td>.11</td>
<td>.27**</td>
<td>.71**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Original hurt</td>
<td>.18</td>
<td>.11</td>
<td>.27**</td>
<td>.71**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Time elapsed</td>
<td>−.04</td>
<td>−.01</td>
<td>.10</td>
<td>−.03</td>
<td>−.07</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>8. Relationship involvement</td>
<td>.06</td>
<td>.16</td>
<td>−.55**</td>
<td>.45**</td>
<td>.60**</td>
<td>−.11</td>
<td>−.17</td>
</tr>
</tbody>
</table>

*p < .001; **p = .07.

and offenders are returned to an even keel, insofar as scores are once again equal. Justice has been found to restore to victims their self-concept (see, for example, Shnabel & Nadler, 2008), their standing within interpersonal relationships (see, for example, Okimoto & Wenzel, 2011), and makes them feel better (e.g. Stillwell et al., 2008). We argued that restoration of these important psychological states provides victims with the confidence to allow themselves to be vulnerable again by forgiving. Accordingly, the just deserts motive mediated relations between punishment and forgiveness in the criminal justice context.

Within interpersonal relationships, meanwhile, it appears that punishment may be indirectly related to forgiveness via the positive influence of the just deserts motive. The effect for just deserts is illuminated by the finding that a competing motive for punishment, revenge, had an inhibiting effect on forgiveness. These results have some important implications for future research. First, it appears that laypeople are able to distinguish between just deserts and revenge motives for punishment. Second, when one considers the effect of punishment on forgiveness, it would seem necessary to take into account what victims expect to achieve by punishing. If they are concerned with making things fair again, it seems that punishment is more likely to result in forgiveness. However, if victims are concerned more with retaliation, then it seems that forgiveness is less likely to occur. Instead, the familiar cycle of revenge and counter-revenge may be more likely to eventuate (e.g. Stillwell et al., 2008).

The effect of just deserts in a criminal justice context may be taken as a given (Carlsmith & Darley, 2008), but future researchers might consider the nuances of punishing for just deserts in close interpersonal relationships. For example, future behaviour control (or deterrence) is considered a ‘happy by-product’ of just deserts (Darley & Pittman, 2003). It is possible that in seeing their transgressor get their just deserts, participants also hoped to send a message that the behaviour will not be tolerated within a relationship that they (the victim) valued and wished to maintain. This interpretation accords with other, related research. For example, Gollwitzer and colleagues (Gollwitzer & Denzler, 2009; Gollwitzer et al., 2011) demonstrated that participants glean greater satisfaction from punishing acts when the offender is led to understand why they have punished rather than when the punishment yields only comparative suffering.

Future researchers may also consider what it is in particular about the restoration of justice that helps people forgive. That is, when and how does justice deliver the different psychological states of improved self-concept, well-being and interpersonal standing that we have posited? Further, they might address the extent to which such valued states operate independently of justice or even in place of justice. In Study 1, for example, we primed punishment, but we should acknowledge that we may also have activated associated constructs such as control and empowerment. And even though we found that punishment may influence forgiveness through the agency of just deserts, it is possible that, if measured, constructs such as control and empowerment could override the effects of justice—thereby replacing justice as the explanatory mechanism. In other words, although we are suggesting that the following model be tested (punishment → justice → valued psychological states → forgiveness), an alternative model is also possible: punishment → valued psychological states → forgiveness. Future researchers should consider measuring constructs such as control and empowerment as alternative explanations for the punishment–forgiveness phenomenon.

Some limitations must be acknowledged. First, in these studies, we have placed an emphasis on contrasting punishment with being unable to punish. Such an approach possesses ecological validity, insofar as victims are not always able to punish. However, we could be criticized for strictly only doing this in Study 1. In Study 3, all participants had punished, and in Study 2, one might argue that the victim never had an opportunity to punish—rather, punishment was met out by a third party. Of course, we note that even when punishment is carried out on the victim’s behalf, the relation between punishment and forgiveness is not altered. That said, context is important. We speculate that when third-party punishment is normative, as it is in the criminal justice system and possibly other contexts such as particular organizational settings, victims may consider third-party punishment to be acceptable and therefore sufficient for forgiveness to occur. However, when first-person punishment is normative—as in intimate relationships—then third-party punishment (e.g. by friends or family members) may be less effective.

Second, it may be argued that it is not clear whether punishment increases forgiveness or whether lack of punishment decreases forgiveness. There was no control condition in Study 1, and in Study 3, punishment itself (which was essentially constant) was directly related to forgiveness only after we controlled for its association with revenge. However, in Study 2, participants in the fail-to-punish condition (where the offender went to court but escaped conviction) behaved the same way as those in the unable-to-punish condition (where the offender was never caught). In the former condition, punishment was possible but did not eventuate; in the latter condition, punishment was not possible. Thus, forgiveness levels were the same regardless of whether the opportunity for punishment was made salient or not. So it appears more
likely that the effect of punishment on forgiveness is driven by the fact that punishment does occur.

Finally, we comment on the broader implications of our findings. First, how do we reconcile them with much previous research showing that inclusive aspects of justice are positively associated with forgiveness and retributive justice is not (for a brief review, see Strelan et al., 2011)? We have no doubt that had we included a restorative justice condition, for example, it would have resulted in even higher levels of forgiveness compared with the retribution (punishment) condition. Thus, we think the present results sit comfortably alongside previous justice-forgiveness research: inclusive aspects of justice facilitate forgiveness more so than retributive aspects of justice, but the latter is more likely to encourage forgiveness than no justice at all.

Second, these results may help to encourage more people to forgive, especially those who are suspicious about forgiving. Theorists have identified several costs or risks of forgiving, including offender recidivism, loss of power and perceptions of the victim as weak (e.g. Baumeister et al., 1998; Lamb & Murphy, 2002). It may be speculated that punishment serves to reduce repeat offending in close relationships or allay fears that forgiveness will encourage the offender to think they can get away with it again. Punishment should also help to restore power relations prior to the conciliatory hand of forgiveness being extended and reduce victim concerns that they will be perceived as weak if they subsequently forgive.

Third, and finally, the results may speak to the different avenues to forgiveness. The vast majority of research on the predictors of forgiveness focuses on what the offender must do to encourage forgiveness (e.g. make amends) and on what victims are doing to re-evaluate their experience to make forgiveness easier to do (e.g. develop empathy for the offender). What has been overlooked is the importance of justice in people’s lives (e.g. Lerner, 1980) and how the getting of justice itself can facilitate forgiveness. Several theorists have already made this point (Exline et al., 2003; Fitness & Peterson, 2008; Tripp et al., 2007; Worthington, 2001), but our package of studies constitute the first systematic empirical attempt to tie together two seemingly incompatible but highly relevant motivations: one, the desire for retribution in response to transgressions (e.g. Carlsmit & Darley, 2008), and the other, a desire for social harmony and the maintenance of valued relationships, often achieved through forgiving (e.g. McCullough, 2008).

ACKNOWLEDGEMENTS

We are grateful for the advice of an anonymous reviewer.

REFERENCES


