An Instrumental Perspective on Apologizing in Bargaining: The Importance of

Forgiveness to Apologize

Abstract

Although very little research in bargaining has addressed how perpetrators should deal with the aftermath of unfair allocations, it has been proposed that an apology may help the reconciliation process. Prior research, however, only focused on whether apologies can reveal positive effects on the reconciliation process but did not focus yet on whether perpetrators are actually willing to apologize. In this paper we investigate perpetrator's willingness to apologize for a trust violation in a bargaining setting. We hypothesized that perpetrators willingness to apologize would be a function of the extent to which the victim of the trust violation is willing to forgive. This effect, however, was expected to emerge only among those perpetrators who are low in dispositional trust. The results from a laboratory study with *actual* transgressions and *actual* apologetic behavior supported our predictions and thus emphasize an instrumental view on apologizing in bargaining situations.

Keywords: Apologies; Perpetrators; Bargaining; Instrumental Perspective; Trust Game; Forgiveness.

An Instrumental Perspective on Apologizing in Bargaining: The importance of

forgiveness to apologize

Bargaining is a breeding ground for unfair allocations. Due to the highly interdependent

nature of bargaining situations, there is a conflict between self-interest and the bargaining

partner's interest (Komorita & Parks, 1996). One important guide that people use to

balance the conflicting interests in these types of bargaining situations is the equality

norm. This norm beholds that all bargaining parties receive an equal share of the

commodity that is to be divided (Van Dijk & De Cremer, 2006). This fairness rule

implies that people do not only care about their own outcomes in bargaining, but also

value the outcomes of others (Blount, 1995; Loewenstein, Thompson, & Bazerman,

1989). Breaking the equality norm is not only considered to be unfair and undesirable

when one receives less than the other party; receiving more is generally considered to be

undesirable too (Blount, 1995; Dana, Cain, & Dawes, 2006; Loewenstein et al., 1989).

People use the equality norm as a guide in bargaining settings. A guide, not only

for their own behavior, but also to base their expectations on of what others will do. In

other words, people expect their bargaining partner to adhere to the equality norm as well

(Van Dijk & De Cremer, 2006). For this reason, violating the equality norm does not only

lead to perceptions of unfairness but also to a decrease in trust (Desmet, De Cremer, &

Van Dijk, 2011). Trust is defined as "a psychological state comprising the intention to

accept vulnerability based upon positive expectations of the intentions or behavior of

another (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395). Based on this definition, it

follows that trust can be violated after deviation from the equality norm because this

deviation violates the positive expectations of the victim that the other party will act in

line with the equality rule. Research indeed suggest that people are aversive towards such an equality violation as people have been shown to make costly choices in order not to violate fairness norms (Dana, Cain, & Dawes, 2006).

After a trust violation, perpetrators can feel motivated to reconcile with the victim. Reconciliation can be valuable to the perpetrator because successful reconciliation leads to a continuation of a cooperative relationship with the victim. Despite the importance of this reconciliation process, research on bargaining has devoted almost no attention to examining the aftermath of unfair offers (De Cremer, 2010). Rather, most studies have examined how trust develops or how it plays a role in maintaining cooperation. As such, hardly any studies – at least to our knowledge - have looked at how violated trust can be repaired. In fact ever since Elangovan and Shapiro (1998, p. 548) noted at the end of the nineties that, "research on the violation of trust has significantly lagged behind interest in the phenomenon of trust", more recent articles have articulated that despite the need to focus on this topic "surprisingly few studies have directly examined how trust may be repaired" (Kim, Dirks, Cooper, & Ferrin, 2006, p. 50). Because trust is considered to be one of the most essential lubricants of our social and economic exchanges (Fukuyama, 1995; Knack & Keefer, 1997; Kramer, 1999), it is important for research to address the kind of actions that are required for reconciliation efforts to succeed.

One important reconciliation tool, available to the perpetrator, is an apology. Apologies address the experienced injustice of the victim (Bottom, Daniels, Gibson, & Murnighan, 2002; Exline, Deshea, & Holeman, 2007; Ohbuchi, Kameda, & Agarie, 1989). An apology is a combined statement of both and admission of wrongdoing and

regret for the violation (Lazare, 2004; Kim, et al., 2009). Apologies directly address the violated positive expectations (i.e. trust) of the victim by implicitly promising that the transgression will not be repeated and thus suggesting that the perpetrator is worthy of being trusted again (Kim et al., 2009).

Research outside the bargaining literature has revealed evidence that relationships can be reconciled more effectively if an apology is given and thereby responsibility for the trust violation is acknowledged (De Cremer & Schouten, 2008; Kim et al., 2004; Lewicki & Bunker, 1996; Tomlinson, Dineen, & Lewicki, 2004). To date, only a few published studies have addressed whether the specific use of apologies has an effect within more economically-based exchanges such as social dilemmas and ultimatum bargaining games (see Bottom, Daniels, Gibson, & Murnighan, 2002; De Cremer, 2010; De Cremer, van Dijk, & Pillutla, 2010; for an interesting field study, see Cohen, 1999). These studies did indeed reveal that apologizing for unfair allocations led to increased cooperation and higher future trust behavior.

What all these studies have in common is that they adopted the perspective of the victim. That is, these studies examined whether and when apologies delivered by the perpetrator have a positive effect on the party suffering from the trust violation. This approach is a first good step towards identifying the important value of apologies in the reconciliation process (De Cremer, Pillutla, & Reinders Folmer, 2011). The most important step, however, is to examine whether or not perpetrators are *willing* to apologize, and when they are most likely to do so. This perpetrator perspective is virtually lacking in the literature and particularly so in the bargaining literature. Consequently, we know very little about whether perpetrators are actually willing to

make use of an apology when resources are allocated in unfair ways. Because reconciliation of relationships is a bi-directional issue, only knowing whether victims desire an apology is of limited value. Therefore, we need to promote our insights into the motives that make perpetrators apologize.

In the present paper, we adopt the perspective of the perpetrator. We study under which conditions perpetrators choose to apologize to the victim. We examine apologizing as a behavior in the context of a modified trust game in which the second party (the one receiving the tripled money send by party 1) violates the fairness norm of equality and thus hampers the trust of the first party. In predicting whether perpetrators would apologize or not, we adopt an *instrumental* perspective, meaning that the choice to apologize by the perpetrator will be motivated by the likelihood that an apology will elicit its intended effect. The effect that we assume that perpetrators strive for when apologizing is to be forgiven by the victim. We consider this approach to be instrumental because the decision to apologize or not becomes conditional on the likelihood of whether the victim will forgive or not.

1. Apologies: An instrumental perspective

We propose that an important reason to apologize is to restore the relationship with the victim, which usually implies that the perpetrator will be *forgiven*. The desire for forgiveness has been identified as an important motive to initiate the reconciliation process (Schnabel & Nadler, 2008). After a transgression, a perpetrator may feel moral inferiority, guilt, or shame. These feelings can lead to an intrinsic motivation to be forgiven by the victim (Schnabel & Nadler, 2008). Indeed, the motive to be forgiven

becomes even more important if the perpetrator wants to continue a cooperative relationship with the victim (Bottom et al., 2002).

It is important to note, however, that while achieving reconciliation may be desirable, perpetrators also take a substantial social risk by apologizing. By apologizing, perpetrators accept blame for their actions (Kim et al., 2009). Hence, perpetrators risk a host of aversive social consequences - like rejection, humiliation and punishment - that would not be the case if they had not accepted blame by apologizing (Exline et al., 2007). An instrumental motivation perspective therefore suggests that perpetrators will be careful to apologize and become strategic when it comes down to apologizing. That is, the decision to apologize by the perpetrator will be conditional on the likelihood that victims are willing to forgive that perpetrator.

Based on this instrumental perspective, it thus stands to reason that perpetrators will be less willing to apologize when the victim seems unforgiving than when the victim seems forgiving. In this case, the decision to apologize should be driven by perceptions of the perpetrator that the victim is indeed willing to restore the relationship. This line of reasoning therefore suggests that when the victim seems forgiving, perpetrators should reason that an apology on their behalf will be instrumental in restoring trust in the relationship. When the victim does not seem forgiving, perpetrators should perceive the delivery of an apology to be less instrumental in achieving reconciliation. As such, expectations that an apology will restore the relationship with the victim should underlie the relationship between perceived forgiveness of the victim and apologizing behavior of the perpetrator.

2. Perpetrator's trust moderating the instrumental approach

Will perpetrators, however, always employ such an instrumental kind of thinking towards apologizing? We argue that whether perpetrators will let their decision to apologize be influenced by the likelihood of whether the victim is motivated to forgive or not, be depend on their level of dispositional trust. The extent to which people differ in dispositional trust is directly related to how people approach interpersonal behaviors in more versus less instrumental ways. We believe there is good reason to expect that particularly low trusters will adjust their apologetic behavior as a function of the perceived likelihood to be forgiven.

One important reason for this hypothesis is that low trusters, in contrast to high trusters, tend to harbor less positive impressions of others, require more reassurance before cooperating, and are less inclined to believe that cooperation will be reciprocated (Tazelaar, Van Lange, & Ouwerkerk, 2004). High trusters are open towards others and more willing to take initial risks. High trusters tend to harbor benign impressions of others, tend to display more immediate cooperation, and are more likely to believe that cooperation will be reciprocated (De Cremer, Snyder, & Dewitte, 2001 ; Tazelaar et al., 2004). As such, high trusters' will initially be more willing to give the benefit of the doubt towards their interaction partner and show more socially risky behavior (Stouten, De Cremer, & van Dijk, 2006). Therefore, high trusters will be more willing to take the risks associated with apologizing while low trusters are less inclined to do this. As such, low trusters can be considered as more strategic in assessing whether an apology will be responded to favorably and will thus apologize more easily if forgiveness is likely to be given (De Cremer et al., 2001; Tazelaar et al., 2004).

To summarize, we predict that low trusters will be particularly influenced in their apologizing behavior by perceptions of the victims' inclination to forgive, while high trusters will be less influenced by the perceived forgiving intentions of the victim (Hypothesis 1). Furthermore, we predict that expectations that an apology will restore the relationship with the victim will mediate the interactive relationship between perceived forgiveness of the victim and perpetrators' dispositional level of trust on apologizing behavior (Hypothesis 2).

3. The Present Research

To test our hypotheses in a controlled manner, we conducted an experimental lab study to investigate actual apologetic behavior. To date, the small number of studies examining the delivery of apologies relied primarily on recall tasks or imagined scenario settings (Meijer, 1998). At least to our knowledge, research has not tested *actual* apologetic behavior. The use of scenarios and free recall tasks is an important first step in understanding apologetic behavior of perpetrator. However, because behavioral intentions and *actual* behavior do not always correspond, it is necessary to test our predictions with respect to actual apologetic behavior (Ajzen & Fishbein, 1977).

Therefore, in our study, we designed a novel paradigm in which participants were induced to commit a transgression against another participant, upon which they were given the opportunity to apologize to the victim. Specifically, to make participants commit a transgression, we modified a standard trust game (Berg, Dickhaut, & McCabe, 1995). The trust game is a coordination game in which players can increase their earnings by trusting each other, at the risk of being exploited. In this game, Player 1 starts with an initial endowment and can decide to transfer any part of his/her endowment to Player 2.

Whatever Player 1 transfers to Player 2 is tripled. Subsequently, Player 2 has to decide

how much of the tripled sum he/she wants to return to Player 1. Thus, the more Player 1

trusts Player 2 to return a fair amount, the more likely Player 1 will be to transfer his/her

endowments. We modified this game to induce unfair behavior by the participant, who

was allocated to the Player 2 position. We did this by creating uncertainty about the

original endowment of player 1; thereby licensing the participant to keep a larger share of

the endowment that player 2 could divide (for the full experimental procedure, see the

Procedure section).

4. Method

4.1 Participants and design.

A total of 153 participants (55 women, 98 men; M(age) = 19.82, SD(age) = 1.59) were

randomly assigned to either the forgiving or unforgiving condition.

4.2 Procedure.

Participants were placed in individual cubicles in front of a computer. Participants'

dispositional trust was measured beforehand, using the 8 item interpersonal trust scale by

Yamagishi (1988; α = .78)[1]. Participants were then asked to engage in a series of tasks

together with another person present in the lab. They would play the previously

mentioned modified version of the trust game. The game was presented as a task on

social decision-making, and all participants were told that they would be Player 2 in the

study. Participants were told that they would be playing an investment game with another

person who was in the lab: Player 1 (all the behavior by Player 1 was preprogrammed). It

was explained that Player 1 had to decide how much of his/her endowment to transfer to

the participant. Participants were clearly told that the amount that Player 1 would send

would be tripled, and that they had to decide which amount to return to Player 1. Subsequently, the task started, and participants learned that Player 1 had transferred 10 chips of significant monetary value, which were tripled to 30 chips. They then had to decide how many of these 30 chips to return to Player 1.

    We modified the game in such a way that participants were likely to make an unfair decision towards Player 1 (i.e., more likely to keep more chips for themselves than to give back to Player 1). We did so by raising uncertainty over Player 1's initial endowment. It was explained that the initial endowment of Player 1 could be anything from 10 to 30 chips; however, the exact endowment was unknown to the participant. Because 10 chips was the lowest endowment possible, we expected that most participants would infer that the original endowment of Player 1 would be larger than 10 chips. To check this assumption, we asked participants to estimate the initial endowment of Player 1 at this point: overall, participants thought the original endowment of Player 1 was 20.41 chips ($SD = 5.57$) large. Because participants estimated the original endowment of player 1 to by larger than 10 chips it meant that Player 1 had chosen not to transfer all his/her chips. From this point of view, we expected that participants would feel justified to keep a larger share of the 30 chips.

    After participants had made their decision how to divide the 30 chips, we revealed that the initial endowment of Player 1 was in fact only 10 chips. Player 1 had thus transferred his/her entire endowment. Participants who had divided the 30 chips unequally (74%) had violated the equality rule and acted unfairly towards Player 1. Participants who had divided the chips equally or returned more than 15 chips (26%) had not committed a transgression. After this feedback concerning the final division of the

chips, we asked participants two questions regarding their perceptions of the fairness of the final division and whether participants thought they violated Player 1's trust. For the group who had not committed a transgression the experiment ended at this point. The majority of the participants, who did commit a transgression, proceeded to the forgiveness manipulation.

4.3 Forgiveness manipulation.

Participants who committed a transgression received a message from Player 1. In the not forgiving condition this message was: "*I have fewer chips than you! I simply do not accept this! I know that I am not the kind of person who forgives this kind of behavior so I will not forgive you this time*". In the forgiving condition, this message was: "*I have fewer chips than you! That is too bad. But I will give you the benefit of the doubt for now. I will forgive you for now but please be cooperative in the future*".

Apology behavior was assessed after this message. Participants were given the choice between two messages to send back to Player 1: "*I want to apologize*" or "*I do not want to apologize*".

Participants' instrumentality perceptions were assessed using three questions (all on a 7-point scale ranging from 1 = not at all to 7 = very much so): (1) To what extent do you think an apology is important for Player 1?, (2) How effective do you think an apology will be to restore your relationship with Player 1?, (3) To what extent do you think an apology will repair the damaged trust between you and Player 1? These items were combined into an average instrumentality perceptions score (α = .85).

We used three items to check our forgiveness manipulation (7-point scale, ranging from 1 = not at all, 7 = very much so): (1) Do you think that Player 1 is somebody who

easily forgives?, (2). Do you think Player 1 is somebody who does not easily forgive? (*Recoded*), and (3). Do you think Player 1 is forgiving? These items were combined into an average forgiveness score (α = .95).

5. Results

5.1 Perceptions of initial endowment size and endowments returned.

To check whether participants indeed overestimated Player 1's initial endowment, we asked participants how many chips they thought Player 1 had originally been endowed with. Overall, participants thought the original endowment of player 1 was 20.41 chips large (SD = 5.57). Participants who committed the transgression thought the original endowment was significantly larger (M = 21.63, SD = 5.00) than participants who did not committed the transgression (M = 16.95, SD = 5.73; $t(151)$ = 4.89, p < .001). In line with these perceptions, participants who transgressed returned less chips (M = 7.88, SD = 3.62) than participants who did not transgress (M = 16.55, SD = 3.31; $t(151)$ = 13.31, p < .001).

5.2 Perceptions of fairness and violated trust.

We modified our trust game in such a way that our participants would commit a transgression. To check whether participants indeed perceived the final division as unfair and as violating trust they were required to respond to three questions. First we asked them *"To what extent do you think the final division is fair?"* (1 = not at all, 7 = completely). A t-test showed that participants who committed the transgression considered the final division to be less fair (M = 2.73, SD = 1.40) than participants who did not committed a transgression (M = 6.00, SD = 1.80, $t(151)$ = 11.749, p < .001). Moreover, a 95% confidence interval of the mean fairness perceptions of the participants

who committed a transgression showed the mean was significantly lower than the scale mean (95% C.I.: 2.47 – 2.99), providing further support that participants who committed the transgression considered the final division to be unfair.

In order to check whether participants regarded the unfair offer as a trust violation we asked them: *"To what extent do you think you violated Player 1's trust?"* (1 = not at all, 7 = completely). Again, participants who committed the transgression thought they violated Player 1's trust significantly more ($M = 5.41$, $SD = 1.22$) than participants who did not committed a transgression ($M = 2.10$, $SD = 1.69$, $t(151) = 13.27$, $p < .001$). A 95% C.I. showed that the mean perception of the trust violation was significantly higher than the scale mean (95% C.I. 5.18 – 5.64), providing evidence that the participants who committed the transgression indeed perceived the unfair division to be a trust violation.

Finally, we checked the extent to which participants thought that player 1 still trusted them *"To what extent do you think Player 1 still trusts you?"* (1 = not at all, 7 = completely). Again, participants who committed the transgression thought Player 1 trusted them less ($M = 3.08$, $SD = 1.50$) than participants who did not commit the transgression ($M = 6.15$, $SD = 1.17$, $t(151) = 11.77$, $p < .001$). A 95% Confidence Interval of the mean trust perception of the participants who committed the transgression again showed that these participants thought that Player 1 was distrustful towards them (95% C.I.: 2.80 – 3.36).

5.3 Forgiveness manipulation check.

A linear regression analysis on our manipulation check scale with the forgiveness manipulation and trust as independent variables revealed a main effect of forgiveness ($\beta = .85$, $t = 16.72$, $p < .001$). No other effects reached significance. Participants indeed

perceived Player 1 to be more forgiving in the forgiving condition ($M = 5.57$, $SD = .79$)

than in the unforgiving condition ($M = 2.40$, $SD = 1.20$).

5.4 Apologetic behavior.

Because our main dependent variable was categorical in nature (either apologize or not

apologize) we used logistic regression to test our hypotheses. Logistic regression uses

odds to test whether a specific response is significantly more likely than chance to be

picked by participants. If an odd is significantly higher than 1, this means (within the

context of this experiment) that it is significantly more likely that an apology is given

instead of no apology. If an odd is significantly smaller than 1, it means that no apology

is significantly more likely to be given than an apology. When an odd does not differ

significantly from 1, it means that it is equally likely that either an apology or no apology

is given (i.e. there is no pattern). Because proportions have more intuitive appeal than

odds, we present, together with the odds, the proportions of apologies given in each

condition. Although these proportions give the same information as the odds, we believe

it helps in interpreting the results.

For our simple effects tests, we followed procedures as outlined by Jaccard

(2001). Simple effects are tested by means of odds ratios. An odds ratio is the ratio

between two odds from two different cells (i.e. conditions). If the odds ratio is

significantly larger or smaller than 1, this means that the odds from those two cells differ

significantly from each other.

We conducted a stepwise logistical regression with the forgiveness manipulation

as a categorical independent variable and trust as a continuous independent variable. The

analysis with the main effects of the forgiveness manipulation and trust in step 1

revealed, first of all, a significant main effect of the forgiveness manipulation (B = -1.74,

Wald = 12.79, $p < .001$) but no main effect of trust (B = .17, Wald = .53, p = .23). In step

2 the main effect of forgiveness remained significant (B = -2.08, Wald = 12.34, $p < .001$),

as was the main effect of trust (B = -1.43, Wald = 3.43, p = .03), but more importantly

and in line with Hypothesis 1, the interaction between forgiveness and trust was

significant (B = 1.19, Wald = 4.46, p = .04; see Table 1).

We compared the forgiveness conditions between high trusters (+1 SD) and low

trusters (-1 SD). Low trusters were more sensitive to the forgiving communication as they

were significantly more likely to apologize when the victim seemed forgiving (proportion

of .96) rather than unforgiving (proportion of .5; B = 3.27, Wald = 11.16, p = .001; odds

ratio = 26.393). High trusters did not differ in their apologetic behavior when the victim

seemed forgiving (proportion of .79) or unforgiving (proportion of .61; B = .87, Wald =

2.06, p = .15; odds ratio = 2.425). The results further showed that when the victim did not

seem forgiving, trust had no significant impact on apologetic behavior (B = .23, Wald =

.59, p = .44; odds ratio = 0.627). When the victim seemed forgiving, low trusters were

significantly more likely than high trusters to apologize (B = -.96, Wald = 4.04, p = .04;

odds ratio = 6.826).

5.5 Instrumentality of an apology.

A regression analysis revealed significant main effects of the forgiveness manipulation

(β = .40, t = 4.72 $p < .001$) and of trust (β = -.17, t = -1.96 p = .03 (one-sided)) on the

perceived instrumentality of an apology. Importantly, we also found the predicted

interaction between perceived forgiveness and trust on perceived instrumentality of the

apology (β = .15, t = -1.77, p = .03 (one-sided)). A simple slopes analysis showed that

trust was a significant predictor when the victim seemed forgiving (β = -.32, t = -2.88, p = .005) but not significant when the victim did not seem forgiving (β = -.02, t = -.12, p = .90).

5.6 Mediation analysis.

To test our second hypothesis, we examined whether perceived instrumentality of an apology mediated the interactive relationship between trust and perceived forgiveness on apologies. Specifically, we expected that perceived instrumentality would mediate the effect of forgiveness information on apologies, but only for those low in trust and not for those high in trust. A bootstrap procedure (Preacher & Hayes, 2008) which tested the simple indirect effect of perceived forgiveness on apologies through perceived instrumentality showed a significant indirect effect for low (-1 SD) trusters (b = .93, $S.E.$ = .30, 95% C.I. -1.44 − -.50) while no significant indirect effect was present for high (+1 SD) trusters (b = -.41, $S.E.$ = .27, 95% C.I. -.87 − .02). This result shows that instrumentality perceptions explained the decision of low trusters to apologize as a function of the likelihood that the victim will forgive, whereas this was not the case for high trusters.

6. Discussion

Dividing valuable resources in an unfair manner can lead to violated trust between the parties involved, which makes future interactions less productive and pleasant. Because unfair allocations easily arise, it is important to understand the mechanisms by which we can repair violated trust in such interdependent settings. Prior research has identified an apology as an effective trust repair strategy (e.g. Bottom et al. 2002; Ohbuchi et al., 1989). However, research to date has neglected to study whether

perpetrators are also willing to actually apologize and if so, when they are most likely to do this. Our present findings show that perpetrators use apologies in a strategic way. That is, they apologize significantly more when the likelihood that the victim will forgive is high. If the likelihood is low, perpetrators are less willing to apologize.

How can these findings be understood? One important reason may be that apologizing entails a considerable social risk. Apologies are often regarded as an acceptance of blame for the transgression, which can give rise to a host of aversive social consequences - like rejection, humiliation and punishment (Kim et al., 2009). This would suggest that it is important for perpetrators to deliver apologies only when they are likely to be met with favorable consequences. One important and favorable consequence in interdependent settings is whether the other party (the victim) will forgive. Forgiveness holds the idea that subsequent interactions will be cooperative and will not include blame of one's prior unfair behavior (McCullough, Fincham, & Tsang, 2003). Thus, apologizing behavior is much more likely to occur if it pays off in terms of being forgiven.

This relationship between anticipated forgiveness and apologetic behavior is further substantiated by the finding that dispositional trust influenced perpetrators' sensitivity to the victim's forgivingness. Trust entails a willingness to be vulnerable to others, and therefore is strongly related to the extent to which people are willing to take social risks (Yamagishi & Yamagishi, 1994). Indeed, we find that low trusters (who do not generally tend to believe people will reciprocate cooperative behavior) are especially sensitive to the perceived forgiveness of the victim, while perceived forgiveness has less impact on decisions to apologize among high trusters, who generally already harbor impressions of benign intent of others. These findings therefore further suggest that

perpetrators' tendency to base decisions to apologize or not on instrumental motives may be rooted in the desire to prevent the social risks associated with apologizing.

It is important to stress that the present findings were obtained by inducing actual transgressions and examining actual apology behavior. This approach – to our knowledge – is the first effort to examine actual apology behavior and transgressions in a controlled bargaining setting. Looking at the actual deliverance of apologies by a perpetrator is important because intentions to apologize may not necessarily correspond with actual apologetic behavior (Ajzen & Fishbein, 1977). We thus urge future research on apologies to examine actual deliveries of apologies as a response to actual transgressions. Our study provides a useful tool to achieve this aim.

Our paradigm succeeded in inducing transgressions with the majority of the participants. By inducing transgressions we needed to rely on deception. Although we are sensitive to the controversies regarding deception in experimental research (for a discussion on the topic, consult Hertwig & Ortmann, 2001), we believe with this type of research it is inevitable to use deception. As participants are motivated to behave in socially appropriate ways, especially when they believe they are being watched, transgressing against another participant in a lab is something that rarely occurs naturally. To create a situation, equal to all participants, under which participants were most likely to transgress, we relied on deception. Because none of the participants expressed any objection to our experimental procedure during the debriefing, we believe our paradigm enables researchers to create real transgressions without being too psychological distressing.

It is important to note that our instrumental hypothesis regarding apologizing includes the notion of reciprocity. As we have mentioned earlier and as our results show, a perpetrator is more willing to apologize when it is likely that this gesture will be reciprocated with forgiveness. A stronger form of this instrumental hypothesis could, however, also be formulated. That is, if it is likely that you are going to be forgiven anyway, perpetrators could also think that there is no reason to apologize anymore. From an economic point of view (i.e. maximizing one's own pay-off) it would indeed make little sense to admit culpability by apologizing if you have already acquired the insurance of your valued good, that is, forgiveness. This 'strong' instrumentality hypothesis can also be considered to be in line with research on moral credentials (Monin & Miller, 2001). Research on moral credentials has shown that once people establish themselves as a moral person, they are more likely to behave in ways that could be interpreted as immoral. If the victim is likely to forgive the perpetrator, the perpetrator can interpret this as an affirmation of his/her morality (Schnabel & Nadler, 2008). Our results contradict this framework, as perpetrators decided to increase their display of moral behavior by apologizing, rather than feeling freed not to apologize (a prediction that would be in line with the moral credentials framework). In line with our findings, research by Wallace and colleagues (Wallace, Exline & Baumeister, 2008) also shows that expressions of forgiveness deter future offences against the victim. Future research should investigate the relationship between moral credentials and expressions of forgiveness in order to further our understanding of both the consequences of showing forgiveness and the regulation of moral behavior.

Future research might zoom in on this interesting question on when our instrumentality hypothesis would be valid and when the strong version of the instrumentality hypothesis would apply more. One possible way of testing this may be to include personality variables that could help us tease apart in what way these instrumentality perspectives influence people's apologetic behavior. Social value orientation predicts whether people approach interpersonal situations as more instrumental versus more social. It could very well be that proselfs deem an apology unnecessary when they interact with a forgiving victim (i.e. they behave accordingly the strong instrumental perspective), while prosocials are more inclined to reciprocate.

Importantly, by focusing on the perspective of perpetrators, our findings also provide a much needed extension to the apology literature, calling previous findings into question. As noted, previous research on apologies has mainly focused on how victims respond to apologies and thus has largely overlooked the perspective of perpetrators. As a result, hardly any research exists examining whether perpetrators actually are willing to deliver apologies, and when they may be likely to do so. Our findings are among the first – at least to our knowledge – to reveal some insights into this question and thus demonstrate the need to also consider the perspective of the perpetrator to arrive at a better understanding of the reconciliation process through the use of apologies.

The present findings contain a hopeful message. Our findings suggest that expressions of forgiveness have the potential to limit a possible downward spiral of unconstructive behaviors that can take place after a transgression. That is, when victims take the initiative to communicate forgiveness, perpetrators are likely to reciprocate by actually apologizing and taking responsibility for their misdeeds. Thus, an initial positive

signal by the victim may elicit the kind of behavior by the perpetrator that is needed to start the reconciliation process. However, there is also a downside to this effect: perpetrators may actually be less willing to apologize when it seems unlikely that apologies will be reciprocated with forgiveness. In other words, if no positive feedback with respect to forgiveness is communicated by the victim, apologies will most likely not be given. This finding challenges the true value of apologies as a trust repair tool. Take, for example, the situation of serious transgressions where victims are likely to be angry, and not very motivated to forgive. Under such circumstances, victims have the strongest need and request for apologies (Exline et al., 2007). However, given the negative reactions on behalf of the victim, our results suggest that perpetrators will be unwilling to apologize.

In conclusion, our findings show that apology behavior by perpetrators is driven by the forgiveness tendencies of victims, thus pointing out the somewhat paradoxical message that the desired response of an apology by the victim actually depends on the positive reaction (i.e. showing forgiveness) of that same victim towards the perpetrator.

7. References

Ajzen, I. & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of the empirical research. *Psychological Bulletin*, *84*, 888-918.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, *10*, 122-142.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, *63*, 131-144.

Bottom, W. P., Gibson, K., Daniels, S. E., & Murnighan, J. K. (2002). When talk is not cheap: Substantive penance and expression of intent in rebuilding cooperation. *Organization Science*, *13*, 497-513.

Cohen, J. R. (1999). Apologies and organizations: Exploring an example from medical practice. *Fordham Urban Law* Journal, *27*, 1447-1482.

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quietly) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*, 193-201.

Desmet, P. T. M., De Cremer, D., & Van Dijk, E. (2011). In money we trust? The use of financial compensations to repair trust in the aftermath of distributive harm. *Organizational Behavior and Human Decision Processes*, *114*, 75-86.

De Cremer, D. (2010). To pay or to apologize? On the psychology of dealing with unfair offers in a dictator game. *Journal of Economic Psychology*, *31*, 843-848.

De Cremer, D. & Schouten, B. C. (2008). When apologies for injustice matter. *European Psychologist*, *13*, 239-247.

De Cremer, D., Snyder, M., & Dewitte, S. (2001). 'The less I trust the less I contribute (or not)?' The effects of trust, accountability and self-monitoring in social dilemmas. *European Journal of Social Psychology*, *31*, 93-107.

Elangovan, A. R. & Shapiro, D. L. (1998). Betrayal of trust in organizations. Academy of Management Review, *23*, 547-566.

Exline, J. J., Deshea, L., & Holeman, V. T. (2007). Is apology worth the risk? Predictors, outcomes and ways to avoid regret. *Journal of Social and Clinical Psychology*, *26*, 479-504.

Fukuyama, F. (1995) *Trust: The Social Virtues and the Creation of Prosperity*, New York: The Free Press.

Hertwig. R. & Ortmann, A. (2001). Experimental practices in economics – A methodological challenge for psychologists. *Behavioral & Bain Sciences* , *24*, 383-451.

Jaccard, J. (2001). *Interaction effects in logistic regression*. Thousand Oaks (CA): Sage.

Kim, P. H., Dirks, K. T., & Cooper, C. D. (2009). The repair of trust: A dynamic bilateral perspective and multilevel conceptualization. *Academy of Management Review*, *34*, 401-422.

Kim, P. H., Dirks, K. T., Cooper, C. D., & Ferrin, D. L. (2006). When more blame is better than less: the implications of internal vs. external attributions for the repair of trust after a compentence- vs. integrity-based trust violation. *Organizational Behavior and Human Decision Processes*, *99*, 49-65.

Kim, P. H., Ferrin, D. L., Cooper, C. D. and Dirks, K. T. (2004). "Removing the shadow
of suspicion: The effects of apology versus denial for repairing competence- versus
integrity-based trust violations". *Journal of Applied Psychology*, *89*, 104-118.

Knack, S. & Keefer, P. (1997). Does social capital have an economic payoff? A cross-
country investigation. *Quarterly Journal of Economics*, *112*, 1251-1288.

Komorita, S.S., & Parks, C.D. (1995). Interpersonal relations: Mixed-motive interaction.
Annual Review of Psychology 46: 183–207.

Kramer, R. (1999). Trust and distrust in organizations: Emerging perspectives, enduring
questions. *Annual Review of Psychology*, *50*, 569-598.

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and
decision making in interpersonal contexts. *Journal of Personality and Social
Psychology*, *57*, 426-441.

McCullough, M. E., Fincham, F. D., & Tsang, J. (2003). Forgiveness, forebearance and
time: The temporal unfolding of transgression-related interpersonal motivations.
Journal of Personality and Social Psychology, *84*, 540-557.

Meijer, A.J. (1998). Apologies: What do we know? *International Journal of Applied
Linguistics*, *8*, 215-231.

Monin, B. & Miller, D. T. (2001). Moral credentials and the expression of prejudice.
Journal of Personality and Social Psychology, *81*, 33-43.

Ohbuchi, K., Kameda M., & Agarie, N. (1989). "Apology as aggression control: Its role
in mediating appraisal of and response to harm". *Journal of Personality and Social
Psychology*, *56*, 219-227.

Preacher, K. J. & Hayes, A. F. (2008). Asymptotic and resampling strategies for

assessing and comparing indirect effects in multiple mediator models. *Behavior*

Research Methods, *40*, 879-891.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after

all: A cross discipline view of trust. *Academy of Management Review*, *23*, 393-404.

Shnabel, N. & Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the

differential emotional needs of victim and perpetrator as a key to promoting

reconciliation. *Journal of Personality and Social Psychology*, *94*, 116-132.

Stouten, J., De Cremer, D., & van Dijk, E. (2006). Violating equality in social dilemmas:

Emotional and retributive reactions as a function of trust, attribution, and honesty.

Personality and Social Psychology Bulletin, *32*, 894-906.

Tazelaar, M. J. A., Van Lange, P. A. M. & Ouwerkerk, J. W. (2004). How to cope with

"noise" in social dilemmas: The benefits of communication. *Journal of Personality*

and Social Psychology, *87*, 845-859.

Van Dijk, E. & De Cremer, D. (2006). Tacit coordination and social dilemmas: On the

importance of self interest and fairness. In D. De Cremer, M. Zeelenberg, & J.

Keith Murnighan (Eds.), *Social Psychology and Economics* (pp. 141-154).

Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Van Lange, P. A. M., De Cremer, D., & Van Dijk, E., & Van Vugt, M. (2007). Self-

interest and beyond: Basic principles of social interaction. In A. W. Kruglanski &

E. T. Higgings (Eds), *Social Psychology: Handbook of Basic Principles,* 2nd

Edition, (pp. 540-561). New York: Guilford.

Wallace, H. M., Exline, J. J., & Baumeister, R. F. (2008). Interpersonal consequences of forgiveness: Does forgiveness deter or encourage repeat offenses? *Journal of Experimental Social Psychology*, *44*, 453-460.

Yamagishi, T. (1988). The Provision of a Sanctioning System in the United States and Japan. *Social Psychology Quarterly*, *51*, 265-271.

Yamagishi T. & Yamagishi M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, *18*, 129-165.

Footnotes

[1]The mean level of dispositional trust did not differ significantly between the participants that committed a transgression (*M* = 3.92, *SD* = .93), and the participants that did not commit a transgression (*M* = 4.02, *SD* = .85, *t*(151) -.62, *p* = .54).

Table 1

Odds and and odds ratio's of an apology per condition

	Trust		Odds ratio for trust within each forgiveness condition
Forgiveness manipulation	Low trust (-1 SD	High trust (+1 SD)	
Forgiving	Odds: 25.918** (Prob: 96%)	Odds: 3.797** (Prob: 79%)	6.826*
Not Forgiving	Odds: .982 (Prob: 50%)	Odds: 1.566 (Prob: 61%)	0.627
Odds ratios for forgiveness within low and high trust	26.393**	2.425	

Note. Proportions that an apology will be given. * = $p < .05$, ** = $p < .01$.