No. 53 june 2010

Urs Fischbacher
Verena Utikal

On the Acceptance of Apologies

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

On the Acceptance of Apologies

10 August 2010*

Urs Fischbacher[1] Verena Utikal[2]

Abstract

An apology is a strong and cheap device to restore social or economic relationships that have been disturbed. In a laboratory experiment we find that harmdoers use apologies in particular if they fear punishment and when their intentions cannot be easily inferred. After offenses with ambiguous intentionality apologizers are punished less often than non-apologizers. Victims expect an apology and punish if they do not receive one. An apology only affects the event of punishment but not the level of punishment. An apology does not help at all after clearly intentionally committed offenses. On the contrary, after such offenses harmdoers do better not to apologize since sending an apology in this situation strongly increases punishment compared to remaining silent.

JEL Classification: C91, D82, D83.

Keywords: Apology, Intentions, Experiment.

---

* We would like to thank Kate Bendrick, Lisa Bruttel, Gerald Eisenkopf, Simeon Schudy, Irenaeus Wolff and the participants at conferences in Kreuzlingen and Copenhagen for helpful comments. Support from the Center of Psychoeconomics at the University of Konstanz is gratefully acknowledged.

[1] Department of Economics, University of Konstanz, 78457 Konstanz, Germany, urs.fischbacher@uni-konstanz.de, Thurgau Institute of Economics, Hauptstrasse 90, CH-8280 Kreuzlingen.

[2] Department of Economics, University of Konstanz, 78457 Konstanz, Germany, verena.utikal@uni-konstanz.de

An apology is the superglue of life. It can repair just about anything.

(Lynn Johnston, Canadian Cartoonist)

1. Introduction

Apologies are remarkable. After an offense they restore social order without amending the offense. No material has been exchanged, yet the relationship between harmdoer and harmed party has improved (Tavuchis (1991)). Apologies can enhance many kinds of economic and social relationships that are upset. They are used in everyday life between individuals, between co-workers and business partners. Apologies can influence the outcome of elections and accelerate peace negotiations. Since people are part of several networks of relations, apologies are omnipresent. But why do apologies work? They cannot undo the offense and they cannot repair the caused damage either. Nevertheless an apology can help to reveal the intention behind the preceding offense. After an offense, a victim usually does not know whether the harmdoer is a friendly person who harmed accidently or whether the harmdoer intended to harm. However, the victim's punishment decision strongly depends on the offender's type. Punishment for intentional offenses is higher than for accidental harm. An apology is an offender's chance to inform the victim about his type. This information can of course be a lie or the truth. Since there is experimental evidence that people have a preference for telling the truth, it is quite feasible that an apologizer is less likely to be an intentional harmdoer than a non-apologizer. Is this the reason why apologies work?

To answer this question we use a laboratory experiment and create an environment where people can economically harm others and where apologies for offenses are appropriate and reasonable. We control for clearly intentional offenses and offenses with ambiguous intentionality which in our design can have two reasons: either they are committed intentionally or due to inability. The novelty of our design is that the offender is always responsible for the offense but did not necessarily commit the offense intentionally. The design also allows the offender to write a message after harming. Our study is the first that not only analyzes the victims' reactions to apologies but also focuses on the offenders' motives for sending an apology. We therefore do not restrict messages to ready-made output but let participants write individual messages. We are interested in the kind of messages offenders will write and whether apologies naturally occur.

We find that an apology is in fact the most common message after an offense. Harmdoers apologize for their offense – even if the apology is costly. However, in contrast to the quote at the beginning, an apology does not 'repair just about anything'. We find that apologies do not 'glue' at all after clearly intentionally committed offenses. On the contrary, after such offenses harmdoers do better not to apologize since sending an apology in this situation strongly increases punishment compared to just remaining silent. In situations where the intention behind the offense is ambiguous, apologies are a very powerful instrument: Harmdoers who apologize are punished less often than harmdoers who remain quiet. Victims seem to expect those responsible for the offense to sincerely apologize. Missing the opportunity to apologize seems to worsen the offense and therefore increases punishment probability. Victims seem to trust that an apology is more than the attempt to get around trouble or punishment. Our results show that this assumption is naïve. We find that offenders primarily apologize if they fear punishment for the offense. Evidently it is not remorse that makes a harmdoer apologize but the hope to prevent punishment. Nevertheless, apologies work. Harmdoers who apologize are punished with lower probability. However, if the apology does not prevent punishment from taking place, it will not mitigate the degree of the punishment either. Our data suggest that people do not partly forgive. They either accept the apology and therefore do not punish, or they do not accept the apology and punish nevertheless.

The rest of this paper is organized as follows. The next session summarizes related literature. Section 3 introduces the experimental design. Section 4 gives predictions. Section 5 presents the results and Section 6 concludes.

2. Related Literature

Our paper is related to studies analyzing behavior after apologies for actually experienced and economically relevant offenses (Ohtsubo and Watanabe (2009), Ho (2007), Skarlicki et al. (2004), Schweitzer et al. (2006), Bottom et al. (2002), Abeler et al. (2009)). Our contribution is the distinction between intentionally committed offenses and offenses with ambiguous intentionality. We are the first to analyze how an apology affects punishment. Our study is also the first one that analyzes offenders' motives for sending an apology.

In psychology there is a large body of evidence that a harmdoer who sends an apology is much more likely to be forgiven than a non-apologizer. Most of these studies can be grouped into three categories. In the first category psychologists present vignettes describing situations in which an offender did or did not apologize. Participants then make judgments about the offender (See for example Ohbuchi and Sato (1994), Weiner et al. (1991), Girard et al. (2002), Ohtsubo and Watanabe (2009), Wada (1998), Scher and Darley (1997)). The second category includes studies where participants have to remember past self-experienced situations. They are told to recall whether the offender did or did not apologize and to give explanations of how they felt in this particular situation and whether they accepted the apology. (See for example Exline et al. (2007), McCullough et al. (1997), McCullough et al. (1998), Schmitt et al. (2004).) The third category uses deceptive role-play with actual offenses (Ohbuchi et al. (1989), Struthers et al. (2008)). All three categories document that apologies have a mitigating effect on anger and increase forgiveness.

Apologies are most effective when they are sufficiently long and come across as sincere (Darby and Schlenker (1989), Shapiro (1991), Skarlicki et al. (2004)), include an expression of responsibility (Scher and Darley (1997), Struthers et al. (2008)), an expression of remorse (Gold and Weiner (2000), Tavuchis (1991), Darby and Schlenker (1989)) and explanations in the form of excuses rather than justifications (Shaw et al. (2003)). It is not clear yet how offers of compensation affect forgiveness. On the one hand, several studies show that compensation payments can increase forgiveness (Bottom et al. (2002), Schmitt et al. (2004), Scher and Darley (1997), Witvliet et al. (2002), Zechmeister and Romero (2002)). On the other hand, Abeler et al. (2009) find that customers who receive an apology instead of monetary compensation forgive significantly more often. The authors conduct a field experiment with three different treatments: An apology treatment, where the customer receives an email including an apology and a high and a low compensation treatment. The authors argue that getting paid money could reduce the intrinsic motivation of customers to forgive (as in Gneezy and Rustichini (2000)) and that an apology might trigger a heuristic to forgive that is hard to overcome rationally. In opposition to this reasoning are the results by Struthers et al. (2008), Skarlicki et al. (2004) and Bennett and Earwaker (1994). They show that an apology does not trigger a heuristic to forgive but that its mitigating effect on punishment crucially depends on the characteristics of the offense. The literature thereby clearly distinguishes between responsibility and intentionality of the harm. In a vignette study Bennett and Earwaker (1994) analyze whether the offender's responsibility affects the acceptance of an apology. They find that the higher the responsibility for the harm, the lower the acceptance of the apology. In Struthers et al. (2008) forgiveness was less likely following an apology when offenders intentionally committed an offense. Skarlicki et al. (2004) present very similar results. In their study receivers of unfair offers in an ultimatum game accept these offers less often after an apology than after no message was sent. Since in this case an unfair offer is always made intentionally, these results show that after intentionally committed harm an apology can backfire and even increase punishment. Those who apologize for an intentionally committed harm may be perceived as self-interested, untrustworthy, and as having an ulterior motive (Fein (1996), Schul et al. (2004)). This might lead to lower acceptance rates. Ohbuchi and Sato (1994) conduct experiments with children and find this effect also with fifth graders. The children accepted a harmdoer's apology only when they believed that the harm was committed unintentionally. Interestingly, second graders were not sensitive to the harmdoer's intent.

Ohtsubo and Watanabe (2009) and Ho (2007) introduce theoretical models predicting that receivers of an apology are sensitive to the cost involved in the apology. Experimental evidence is ambiguous. Ohtsubo and Watanabe (2009) find their hypotheses confirmed. In their experiment participants in the costly apology condition abstained from sending a complaint message to their offender after an apology. In the repeated trust game by Ho (2007), costs do not affect the event of forgiveness, which is measured by the amount entrusted in the following period.

An apology's effect on reputation has been analyzed in Schweitzer et al. (2006) and Bottom et al. (2002). Schweitzer et al. (2006) use a trust game in order to test how apologies can repair trust after an offense within a repeated interaction. Schweitzer et al. (2006) find that an apology alone does not facilitate trust recovery. The apology has to come along with a promise for future trustworthy behavior. Bottom et al. (2002) conduct a prisoner's dilemma and find that apologies indicating good intentions for the future have a positive effect on trust recovery in repeated interactions.

Apologies also differ with respect to culture and gender. Asians apologize more than Americans (Takaku et al. (2001)) and women apologize more than men (Tavuchis (1991), page 127). Frantz and Bennigson (2005) find that apologies expressed at a later stage of a conflict are more effective than earlier ones, and that this effect is mediated by feeling heard and understood.

In law there is a broad literature on apologies, too. Here the main question of interest is whether an apology is a possible mechanism to avoid a law suit. For overviews see Cohen (2002) or White (2009).

3. Experimental Design and Procedure

Our basic design is similar to Ho (2007) who uses a trust game with an apology option at the end. We use a sequential prisoner's dilemma with apology option and punishment. Additionally we manipulate the mechanism how to cooperate. For cooperation, subjects have to correctly answer a question. Defection results, when they answer the question incorrectly. This feature allows that defection can be the result of inability or of intentional unkindness. In detail, the sequence is the following:

1. At the beginning of the game both players receive an endowment of 60 points in order to avoid negative payoffs throughout the game.
2. Player A then receives a multiple choice question.
   If he gives the correct answer, he loses 40 points and his partner player B receives 120 points. If he answers incorrectly, points do not change.
3. Both players learn about player A's result. Next, player B receives a multiple choice question, too. If he answers it correctly, he loses 40 points and player A receives 120 points. If he answers incorrectly, points do not change.
   Therefore, if both players answer their questions correctly, they receive 80 points each. However, player B maximizes his payoff by giving a wrong answer after a correct answer by player A.
4. The players learn whether the answers were correct and player B can send an individual message to player A. We are interested in whether offenders use the message to apologize for their harming.
5. Next, player A can deduct points from player B. One deducted point costs player A 0.2 points. Punishment is restricted such that it cannot yield negative payoffs.

The questions were easy but not trivial. For example, we asked for the capital of Japan, giving Tokyo, Osaka, Yokohama and Kyoto as possible answers. By ensuring that not all participants could perfectly solve the questions, we induced uncertainty of intentions. Not giving the right answer to the question can be chosen intentionally[3] or due to inability. We provide a measure for the question's difficulty and an opportunity for the players to form an individual view on the difficulty of their partners' questions. For that purpose the players receive their partners' multiple choice questions at the same time. If they answer this solo question correctly, they receive 5 points. We call the fraction of players who are able to solve their partner's multiple choice question the *solvability benchmark*. On average participants could solve 85% of the questions.[4]

We run five treatments. The first treatment is the control treatment (*baseline)* as explained above. In the second treatment *no punishment* no punishment is possible. We use this treatment to analyze whether offenders still apologize when they do not fear punishment. Third, in the treatment *costly* writing a message costs 5 points. The results of this treatment will show whether the decision to apologize depends on costs and whether costly apologies

[3] In a laboratory experiment Utikal and Fischbacher (2009) analyze how people attribute intent to helpful and harmful actions. They find that when the harming agent is economically strong, people perceive the offense as intentional. Therefore we define one possible reason for giving a wrong answer as intent.
[4] We discuss the implications in the results section.

are more credible. In the fourth treatment *no apology option,* player B cannot send messages. We thereby control how people cooperate and punish deviators when no apology is possible. The fifth treatment *no quiz* varies the potential assignment of intentions. In the quiz games a wrong answer can be due to intentional harm or due to inability. In *no quiz*, an offense is always caused intentionally. In this treatment we use the same parameters as before. However, participants do not have to answer questions. They just decide whether they want to give up 40 points in order to cede 120 points to their partner. Table 1 gives an overview of the treatments.

	Baseline 3 sessions (78 subjects)	No punishment 2 sessions (50 subjects)	Costly 3 sessions (74 subjects)	No apology option 2 sessions (52 subjects)	No quiz 2 sessions (50 subjects)
Quiz	X	X	X	X	
Apology option	X	X	X		X
Apology costless	X	X			X
Punishment	X		X	X	X

Table 1: Treatments

The whole procedure was common knowledge. We conducted 12 sessions in the time from June to November 2009. All sessions were conducted at the LakeLab (TWI/University of Konstanz) with a total number of 356 participants. Before the experiment started, subjects were randomly assigned to their role as player A or B. Players kept their role throughout the game in order to avoid learning the content of others' messages. The experiment lasted 10 rounds. We used a perfect stranger matching in order to avoid repetition effects and to keep players A from receiving identical messages. Participants received the income of all periods. One point translated into 0.01 euros. The experiment took about 60 minutes, average income of participants was 10.93 euros (14.87$) plus a show-up fee of 2 euros (2.72$). The games were programmed with z-Tree (Fischbacher (2007)). We recruited participants using the online recruiting system ORSEE (Greiner (2004)). Each subject sat at a randomly assigned PC terminal and was given a copy of instructions.[5] A set of control questions was provided to ensure the understanding of the game. The experiment did not start until all subjects had answered all questions correctly. We ensured that no subject participated more than once in our experiment. We rule out spillover effects across and within sessions by giving every player a different question.

[5] A translation of the instructions can be found in the appendix.

4. Predictions

What kind of information do people confer when they apologize? The typical apology contains a statement of remorse and a statement that the outcome will not occur again in the future. A special case of the latter statement is the claim that the outcome was not caused intentionally. In this section, we investigate how and why an apology can work. First, we present the formalization of the game that players are playing in our experiment. In the second part we present theoretical elements that can explain why apologies work, complemented with the most important experimental evidence. We integrate these elements into one theoretical model. Finally, we characterize the model's theoretical predictions.

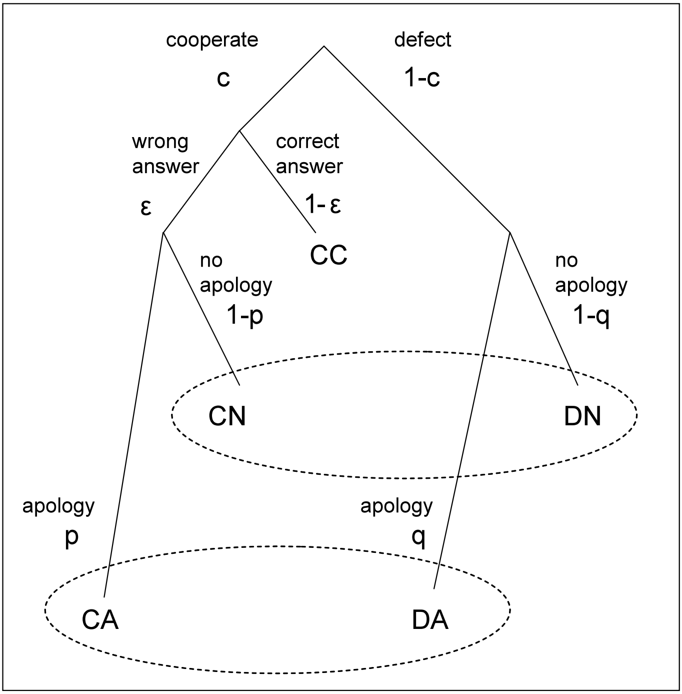

Figure 1: Decision set of player B after player A has correctly answered the question

Figure 1 shows the subgame starting at Player B's move after player A has correctly answered the question. Note that in contrast to the game theoretical outcome based meaning of the expressions "defection" and "cooperation" in the prisoner's dilemma, we define them intention-based. We distinguish between players who want to answer incorrectly (defectors) and players who want to answer correctly (cooperators). This way, also players who answered erroneously incorrectly are cooperators.

The game tree describes that player B either tries to give the correct answer, i.e., he cooperates (c) or intentionally gives the wrong answer ($1-c$), which means that he defects. If he cooperates, he will be able to give the correct answer with probability $1-\varepsilon$. If he defects, we assume that player B is able to give the wrong answer with certainty. This is plausible because the answers were easy and it was even easier to find at least one wrong answer. After intentionally or unintentionally giving the wrong answer, player B can apologize, which means he claims that he wanted to give the correct answer. We will show that if people are averse to lying, those who apologize are more likely to have wanted to give the right answer than non-apologizers. Therefore, they care more about the other players' payoff and deserve less punishment.

Of course, the information of whether player B simply does not know the answer or answers intentionally incorrectly is private. When receiving an apology after a wrong answer by player B, player A does not know whether player B cooperated or defected. Therefore, an apology after cooperation (*CA*) and an apology after defection (*DA*) belong to the same information set of player A. No apology after cooperation (*CN*) and no apology after defection (*DN*) form another information set. We define p as the proportion of cooperators who apologize and q as the percentage of defectors who apologize.

First, we assume selfish preferences and derive the subgame perfect Nash equilibrium using backward induction. Regardless of what happens, player A does not punish. Therefore, player B defects and does not solve the task. Anticipating this player A defects too. Since the players' behavior does not depend on the preceding behavior of the other player, text messages cannot have an impact on behavior – neither on punishment nor on the cooperative behavior in the knowledge questions.

Theories that model non-selfish motives based on outcome-oriented preferences like Fehr and Schmidt (1999) or Bolton and Ockenfels (2000) also predict that apologies have no impact on behavior. These models assume that people have an additional component in the utility function, which captures disutility from inequity. In these models, utility depends not only on the own payoff but also on the payoff of the other players. However, utility depends only on the final allocation, i.e. it does not depend on the procedure. This implies that punishment should only depend on the outcome of the decision, in particular it does not depend on the existence or content of the message and, even when known, it would not depend on the reason why a player B answered incorrectly.

However, in this paper we hypothesize that apologies truthfully transmit regret concerning an unintentional outcome. Thus, the first two relevant elements for understanding apologies are that people are lying averse and also liar-averse. That means that people have a preference for truth-telling - for themselves and also with respect to others. These assumptions are supported by Hurkens and Kartik (2009), Charness and Dufwenberg (2006), Lundquist et al. (2009), Sánchez-Pagés and Vorsatz (2007) and Sánchez-Pagés and Vorsatz (2009). The third important assumption is that people care about intentions. Several experiments have shown the importance of intentions (Blount (1995), Charness and Levine (2007), Brandts and Sola (2001), Falk et al. (2003), Falk et al. (2008)) and in response to this evidence, theories have been developed that take intentions into account. Two main approaches in the modeling of intentions can be distinguished. In theories based on Rabin (1993), intentions are inferred with respect to which choice was made in comparison to the possible alternatives. In such a model, a purposeful wrong answer by player B would be considered as intentionally unkind because in expected terms, it was the most harmful action player B could take. On the other hand, an erroneously wrong action was not planned to harm player A and was the best action that was available to B. However, the intention cannot be observed. When B wants to cooperate but fails the same outcome results as when B intentionally gives the wrong answer.

Levine (1998) takes a different approach to model intentions but we will show that also in his model, planned wrong answers are considered as less kind than unplanned wrong answers. In the model of Levine, players differ in their concern for the other players' payoffs. The value of this concern for player *i* is defined as $\alpha_i \in (-1,1)$.[6] Furthermore, ceteris paribus, players reward those with a high positive other regarding concern and punish those with a low, negative other regarding concern. This reciprocity parameter is denoted as $\lambda \in [0,1)$. The model of Levine values the other players' payoffs with a weight of $\kappa = (\alpha_i + \lambda \alpha_j)/(1+\lambda)$. Since the Levine model allows for player types we will use his model to incorporate intentions and complement it with lying and liar aversion in the following way. First, if people lie themselves, they experience a disutility $\delta_i \geq 0$. Second, if they believe that the other person lies with probability r_{ij}, the weight of the other player's payoff is reduced to $\kappa = (\alpha_i + \lambda \alpha_j - \eta_i r_{ij})/(1+\lambda)$, where $\eta_i \in [0,1]$ is the liar aversion of player *i*. In the following, we analyze the game based on this model. Here apologies can have an impact when players have preferences as assumed in the extended Levine model.

[6] This restriction is assumed by Levine (1998). However, it is irrelevant for our results.

Note that pooling equilibria in which all players apologize or no player apologizes cannot be excluded in the general case – even when we additionally restrict equilibria to satisfy the intuitive criterion. Furthermore, we note that multiple equilibria can exist. Nevertheless, some general characteristics of the equilibria can be shown. Let us focus on those players B who answered the question incorrectly in the case when apologies matter, i.e. in a separating equilibrium in which punishment differs between those who apologize and those who do not. We will show that in this case those who apologize for their offense have on average a higher value of α_i than those who do not apologize. Therefore they deserve less punishment, which justifies that they apologize. An apology is credible when lying aversion is sufficiently large to prevent some defectors from apologizing. Such a separating equilibrium does not always exist. Nevertheless, the following proposition holds.

Proposition Assume that players are risk neutral and have a utility function as in the modified Levine model outlined above. As in the Levine model, we assume that the reciprocity parameter λ is common knowledge. Further, we assume that the other-regarding concern α_i, the lying aversion δ_i and the liar aversion η_i are private knowledge. Finally, we assume that the distributions are continuous, common knowledge, and that α_i, δ_i, and η_i are independent. Furthermore, $\varepsilon < 1$. Then for any perfect equilibrium the following properties hold.

a) In the quiz games ($\varepsilon > 0$) with punishment and apology option, there is either a pooling equilibrium or punishment is at most as high after an apology than after no apology, i.e. $pun_A \le pun_N$.

b) Assume $pun_A < pun_N$. If player B cooperates and was nevertheless unable to answer correctly, he will apologize, i.e. $p = 1$.

c) Assume $pun_A < pun_N$. On average, players who apologize have an α_i that is at least as high as that of players who do not apologize.

d) When all players have a strictly positive lying aversion δ_i, then at least as many players apologize in the punishment condition as in the no punishment condition.

e) If $\varepsilon = 0$ (*no quiz* treatment), punishment is at least as high after an apology than after no apology. Thus $pun_A \ge pun_N$.

Proof of the Proposition

a) If punishment was higher after an apology, no one would apologize and a pooling equilibrium results.

b) If punishment is lower after an apology, all players apologize when apologizing is sufficiently cheap. Since it is costless for those who cooperate, they apologize.

c) Let δ^* be the critical value that makes defecting players indifferent between apologizing and not apologizing. First, we characterize the players who defect and do not apologize. These players have a δ_i^H with $\delta_i^H > \Delta$, where Δ is the difference in payoff after apologizing and not apologizing. Since cooperation always results in a higher payoff for player A than defection, players who defect have a lower κ_i than those who cooperate. To be precise, there is a critical value κ^* below which players with $\delta_i^H > \Delta$ defect. For $\kappa_i<\kappa^*$, it is

$$U(DN,\delta_i^H) > U(C,\delta_i^H) \quad (1)$$

Players who defect and apologize have a δ_i^L with $\delta_i^L \leq \Delta$. Since lying aversion is only relevant after defection, it follows

$$U(C,\delta_i^L) = U(C,\delta_i^H) \quad (2)$$

$$U(DN,\delta_i^H) = U(DN,\delta^*) \quad (3)$$

$$U(DA,\delta^*) = U(DN,\delta^*) \quad (4)$$

$$U(DA,\delta^*) < U(DA,\delta_i^L) \quad (5)$$

$$\Rightarrow U(DA,\delta_i^L) > U(C,\delta_i^L) \quad (6)$$

This means that players with $\kappa_i<\kappa^*$ and $\delta_i^L \leq \Delta$ defect and apologize. Thus, the critical κ^{**} for all players with $\delta_i^L \leq \Delta$ is higher or equal to κ^*, i.e. $\kappa^*\leq\kappa^{**}$. Thus, the group of players who do not apologize consists of all players with $\delta_i > \Delta$ and $\kappa_i<\kappa^*$, while the group of players who apologize consists of all players with $\delta_i \leq \Delta$ and $\kappa_i<\kappa^*$, as well as players with $\kappa_i>\kappa^*$. These players are the players with $\kappa^*<\kappa_i<\kappa^{**}$, and those who cooperate but fail to answer correctly. Because α_i and δ_i are independent, this implies that the average κ_i of the players who do not apologize is lower or equal to the average κ_i of the players who apologize. Finally, because α_j, and r_{ij} are in equilibrium independent of the types and the parameters α_j and η_j are independent, the lower average κ_i translates into a lower average of α_i.

d) No type of player defects when the punishment option exists and cooperates when the punishment option does not exist. Therefore, the share of cooperators is (weakly) higher in the punishment condition than in the no punishment condition. Furthermore, all cooperators in the punishment condition apologize and no defector in the no punishment condition apologizes because they have strictly positive lying cost δ_i.

e) Apologies after clearly intentionally committed offenses are lies. If somebody claims that he did not commit the offense intentionally, he is truly dishonest. Assuming people are liar averse, they will punish such a behavior. q.e.d.

We cannot make a prediction concerning the effect of the apology option. Intuitively, i.e. ignoring changes in cooperation in response to the apology option, we expect that the punishment levels of players who do not have the opportunity to apologize are between the punishment levels of players after having received an apology and after no apology.

Similarly, there is no unambiguous effect when apologies are costly. Introducing a marginal cost suggests that writing an apology becomes more demanding, fewer defectors as well as cooperators apologize. We summarize the preceding reasoning to the following predictions.

Prediction 1 Harmdoers who face possible punishment are more likely to apologize than harmdoers who do not face possible punishment.

Prediction 2 In the quiz games apologizers will be punished less than non-apologizers.

Prediction 3 Offenders without apology option will be punished more than apologizers and less than non-apologizers.

Prediction 4 In the no quiz game punishment is higher after an apology than after no apology.

5. Results

This section is divided into three parts. First, we turn towards behavior in the sequential prisoner's dilemma. Second we present motives for apologies. Third, we focus on the effects of apologies on punishment. In the latter two sections we focus on behavior after offenses only, namely the situation when player B did not answer correctly although player A did. In this situation apologies can occur.

Results on the Prisoner's Dilemma

In this section, we present the results of the prisoner's dilemma. We first focus on the results of the quiz game. In the quiz game treatments participants can cooperate by answering a question. Player A's performance can be interpreted as trust in a correct answer of player B. We find that possible punishment significantly improves player A's performance. Player A's performance does not differ from the baseline when messages are costly. When participants cannot write apologies player A's performance decreases.

Player B receives the same question as player A in order to create a solvability benchmark. 86% of players B are able to correctly answer the question of player A and 83% of players A give the right answer to the question of player B. This difference is not significant which means that the questions have the same level of difficulty.[7] As Figure 2 shows, player A's performance is significantly below this solvability benchmark in all treatments. The results of the related regression are presented in Table 4 (column 1).

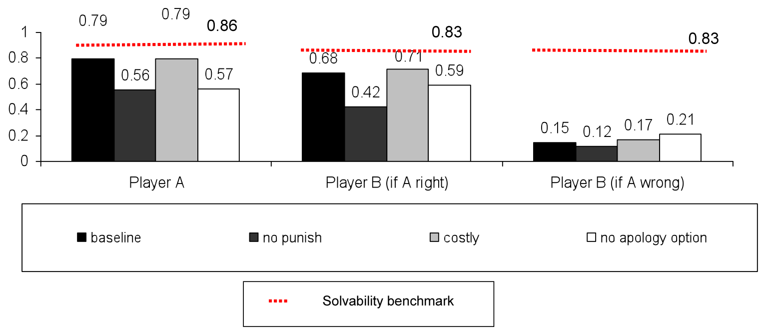

Figure 2: Performance in quiz games

Let us now turn to player B. Figure 2 indicates that player B can find himself in two different situations. Player A might or might not have answered the question correctly. If player A answered correctly, a following correct answer by Player B can be interpreted as trustworthiness. We find that player B's performance is always below the solvability benchmark. When A answered correctly in the baseline treatment 68% of players B answer correctly, too. Without the punishment threat player B's performance significantly drops down to 42%. Performance in the costly message treatment is 71% and does not vary from the baseline treatment. When participants do not have the option to apologize, performance significantly decreases to 59%.[8] The results of the related regression are presented in Table 4 (column 2).

When A did not answer correctly 15% of players B answer correctly in the baseline treatment. When no apology is possible this share slightly increases. We do not find that the other treatments have an effect on performance. Player B's performance strongly depends on

[7] Probit regression with standard errors clustered on session: $p>0.1$

[8] This result is in line with findings by Ellingsen and Johannesson (2008) who show that cooperation increases with anticipated feedback.

player A's behavior. This means that a correct answer by player A significantly increases the share of correct answers by player B.[9] Conditional cooperation is therefore very strong. Table 3 (column 3) shows the results of the corresponding regressions.

We do not find that behavior in the prisoner's dilemma depends on the way the dilemma is implemented. There are no significant differences between the quiz games and the no quiz game..[10]

Motives for apologies

		Treatments			
Category	**Example**	**Baseline**	**no punishment**	**costly**	**no quiz**
Frequency of offense		0.25	0.32	0.23	0.18
Number of offenses		98	80	84	44
No message		0.27	0.76	0.51	0.52
Apology: **Admission of blame-worthiness and regret**	I am sorry. I thought the answer was…	0.33	0.09	0.28	0.11
Apology & other		0.08		0.03	0.04
Admission of blame-worthiness (without admission of regret)	I thought the answer was…	0.05			0.05
Blameworthiness & other		0.02			0.09
Admission of regret (without admission of blameworthiness)	I am sorry.	0.07	0.08	0.04	0.09
Other		0.17	0.08	0.14	0.09

Table 2: Percentage of message categories after offense with apology option

In all treatments except in the treatment *no apology option,* player B was endowed with an individual message option. In order to analyze the written messages 19 additional subjects were recruited to independently sort messages into at least one of the 5 following categories: *admission of blameworthiness, admission of regret, request for mercy* (Please do not punish me), *small talk* (jokes, etc), and *invitation to punish* (Please punish me). In case of disunity between the raters, we applied the majority rule. Messages were only sorted into one particular category when the majority of raters decided to do so[11]. For the analysis we define an *apology* according to Schlenker and Darby (1981) as an 'admission of blameworthiness

[9] Probit regression with standard errors clustered on session: $p<0.01$

[10] The only small difference is that slightly fewer players B cooperate after a non-cooperative move of player A in the no quiz game (Probit regression with standard errors clustered on session: $p<0.1$) This result might be driven by participants who simply dislike giving a wrong answer and gain some utility from answering correctly.

[11] Following this procedure two messages were taken out of the sample because there was no majority for any category.

and regret for an undesirable event'.[12] Due to a small number of observations and for reasons of simplicity we combine the categories *request for mercy*, *small talk*, and *invitation to punish* into the category *other*. Table 2 presents the categories' relative frequencies.

Against expectations only few harmdoers use the message to explicitly emphasize their good intentions. On average only 5% use the expressions 'unintentional' or 'not intentional". A much more common approach for harmdoers is to give a statement of blameworthiness indicating why they did not answer correctly. This kind of message can be interpreted as an implicit expression of unintentionality.. Although some offenders write pure admissions of blameworthiness and pure admissions of regret, the most frequent message after an offense is in fact an apology.[13] Therefore we will focus on the motives and effects of this kind of message

Result 1 Harmdoers who face possible punishment are more likely to apologize.

Table 4 (column 4) shows that the harmdoers' decision of whether to apologize depends on two variables: First, possible punishment significantly increases sent apologies. In the baseline treatment 41% of offenses are followed by an apology. Without the punishment 9% of offenses are followed by an apology. This confirms Prediction 1.

	Correct answer of player A in quiz games	Correct answer of player B after correct answer by A in quiz games	Correct answer of player B after wrong answer by A in quiz games	Apology after an offense in games with apology option
	(1)	(2)	(3)	(4)
No punishment	-.085***(.028)	-.145***(.041)	-.055(.045)	-.060***(.010)
Costly	-.023(.025)	.009(.030)	.030(.048)	-.032***(.012)
No apology option	-.136*** (.046)	.058**(.027)	.147***(.025)	
No quiz				-.067***(.013)
Solo question	.097***(.035)	-.229***(.028)	.147***(.034)	
No of observations	3560	2586	974	1235
Wald χ^2	23.67	358.33	880.09	94.50
Prob>χ^2	.0001	.0000	.0000	.0000
Pseudo R^2	.0286	.0859	.3538	.0411
Number of clusters	10	10	10	10

*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 4: Probit regressions, reporting marginal effects, with robust standard errors, standard errors in parentheses clustered on session

Result 2 Harmdoers who face apology costs are less likely to apologize.

[12] Goffman (1971) states that an apology contains a promise of more acceptable behaviour in the future. Since we have a perfect stranger matching, this definition does not apply to our environment.

[13] Apologies evolve almost exclusively after offenses. There is only one exception: When both players answered their questions incorrectly in the quiz games, in the baseline treatment 4% of players B apologize for their fault. This behavior can for example be explained by fear of punishment. A more general explanation would be that certain people just dislike giving wrong answers or letting the other player down, feel guilty after a failure and apologize.

In the *costly* treatment messages cost 5 points. When messages are costly the frequency of apologies significantly decreases from 41% to 30%.

Result 3 Harmdoers who have assignable negative intentions are less likely to apologize.

Assignable intentions decrease the number of sent apologies significantly. In the quiz game offenses can be caused either intentionally or due to inability. In the no quiz game an offense is always caused intentionally. We find that apologies are present after only 10% of after intentionally committed offenses.

Effects of apologies on punishment

In the quiz games the receiver of an apology faces uncertainty. He cannot be sure whether the harmdoer committed the offense intentionally or due to inability. He also does not know whether the harmdoer sent the apology due to honest regret for the offense or in order to avoid punishment. We are interested in how receivers of apologetic messages react to this uncertainty. We also determine the benchmark for punishment after an offense without apology option. Doing this, we can investigate whether apologizing decreases punishment or whether not apologizing increases punishment. Eventually, we focus on victims reactions to apologies after clearly intentionally committed offenses.

Result 4 After an offense with ambiguous intentionality apologizers are punished less often than offenders who remain silent.

Figure 3 shows the fraction of players A willing to punish harmdoers in the four treatments with punishment option. We distinguish between punishment probability after *no message* and after an *apology*. First we consider the quiz game treatments with uncertain intentionality and message option. In *baseline* the punishment probability after an apology is 17 percentage points lower. In *costly* the punishment probability after an apology is 21 percentage points lower. We conclude that after an offense with ambiguous intentionality apologizers are punished less often than offenders who remain silent. This confirms Prediction 2.

No other message yields the same effect. We find that admissions of regret or blameworthiness or messages including other content do not have a significant effect on punishment. That means it is not just any message that mitigates punishment after an offense, but it has to be an apology. We do not find that an apology costs have an effect on punishment. Costly apologies do not seem to be not more credible than costless apologies.

Table 4 (column 1) presents the results of the corresponding regressions.

Result 5 Apologizing does not decrease punishment probability. Not apologizing increases punishment probability.

We ran the additional treatment *no apology option* in order to determine whether an apology decreases punishment or whether no apology increases punishment. 52% of participants punish harmdoers in the *no apology option* treatment.[14] We define this fraction as the punishment benchmark. As Table 4 (column 4) shows, the punishment probability after an apology does not differ significantly from this benchmark. This means apologizers are not punished less than offenders who did not have the option to apologize. However, the punishment probability after no message was sent although an apology was possible is significantly higher than the benchmark. This result suggests that victims have a demand for apologies if they are possible. Not sending an apology despite being able to do so increases punishment probability compared to situations where apologies are not permitted. This partly confirms Prediction 3.

Result 6 An apology affects the event of punishment but not the level of punishment.

Until now, we focused on punishment probabilities. Our design also allows us to measure the effects of apologies on punishment points. Figure 4 shows the average of punishment points assigned to player B by player A. We find that after offenses with ambiguous intentionality, apologizers (48% in *baseline*, 58% in *costly*) are punished less than non-apologizers (65% in *baseline*, 79% in *costly*). Table 4 (column 5) presents the results of the regression. However, if we control for conditional punishment only, the punishment-decreasing effect of an apology vanishes. If harmdoers apologize after an offense with uncertain intentionality, the probability for punishment decreases. However, if the apology does not prevent punishment, it will not mitigate punishment either. (See Table 4 (column 6).)

Result 7 After a clearly intentionally committed offense punishment probability and punishment level after an apology are higher than after silence.

Now we turn to the *no quiz* treatment where offenses are clearly intentional. 57% of harmed players A decide to punish when their counterpart remains silent. Offenders who remain silent receive 30.21 punishment points on average. After an apology, 71% are willing to punish. Apologizers are punished 45.43 points on average. We find that after a clearly

[14] Since in *no apology* no message is possible, there is no distinction between *no message* and *apology*.

intentional offense, apologizers are punished significantly more often and significantly more than non-apologizers. This confirms Prediction 4.

We find that punishment and apology behavior are perfectly aligned. After clearly intentionally committed offenses apologizers are punished more often. Harmdoers apparently foresee this and apologize less. After offenses with uncertain intentionality apologizers are punished less than non-apologizers. Harmdoers react accordingly and apologize.

The optimal strategy for player A is to cooperate. The optimal strategy for player B in the quiz treatments is to answer his multiple choice question incorrectly and to apologize for it. In the quiz treatment player B should not cooperate and not apologize.

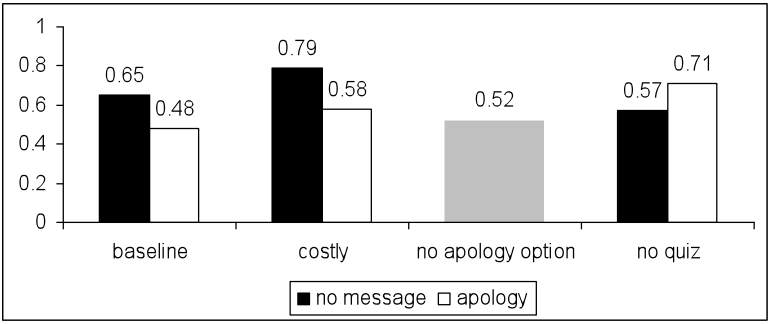

Figure 4: Punishment probability after an offense with punishment option

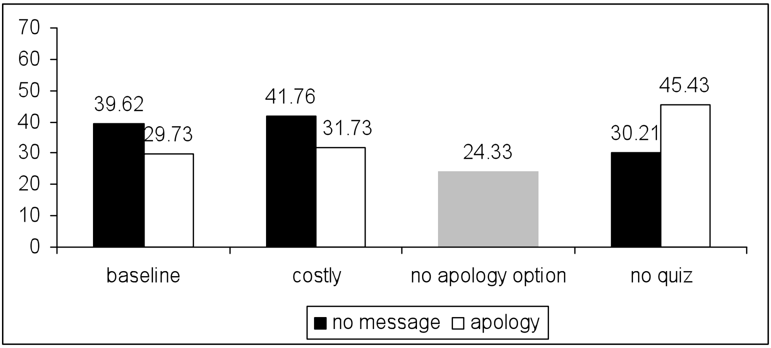

Figure 5: Punishment points for player B by player A after an offense with punishment option

	Punishment probability after offense with punishment option				Punishment points after offense with punishment option	
	quiz games with apology option	no quiz game	games with apology option	quiz games	games with apology option	games with apology option, punishment>0
	(1)	(2)	(3)	(4)	(5)	
Apology	-.201**(.078)	.136(.130)	-.202***(.077)	-.001(.096)	-23.65*(12.59)	5.97(12.57)
Regret	-.111(.133)	.165***(.054)	-.111(.130)		-7.95(23.71)	6.80(18.46)
Blameworthiness	-.107(.161)	.244**(.107)	-.107 (.157)		-22.30(126.00)	-13.34*(7.28)
Other message	.008(.149)	.165(.163)	.008 (.145)		16.18(26.64)	13.58(11.96)
Message costly	.107(.091)		.105(.086)		7.28(26.64)	-9.32(21.03)
No quiz			-.111(.125)		-22.47 (26.64)	-5.77(7.69)
Apology x No quiz			.263**(.068)		65.47(231.30)	14.68(11.94)
Regret x No quiz			.235** (.063)		49.69(35.34)	11.92(18.62)
Blame x No quiz			.286**(.081)		118.70(310.49)	70.07(9.41)
Other x No quiz			.160(.156)		92.81(241.45)	304.24***(40.23)
No message x Apology option				.219**(.099)		
Constant					28.81(19.29)	67.04***(6.96)
No of observations	180	44	224	194	224	144
Wald χ^2				11.31	324.00	2.72e+06
Prob>χ^2				.0035	.0000	0.000
Pseudo R^2	0.0435	.0388	.05428	.0346	.0115	0.0150
Number of clusters	6	2	8	8	8	8
Left-censored					80	0
Uncensored					95	95
Right-censored					49	49
Replications					214	106

*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 4: Probit (reporting marginal effects) and Tobit regressions[15] with robust standard errors, standard errors in parentheses clustered on session, bootstrapped standard errors for Tobit regression

[15] Since punishment is restricted to 0 ≤punishment≤100 we use a Tobit regression.

6. Conclusion

An apology, no matter how sincere or affective, does not and cannot undo what has been done. And yet, in a mysterious way and according to its own logic, this is precisely what it manages to do. (Nicholas Tavuchis (1991)). This paper partly discloses the mysterious way and logic of the role of an apology. We shed light on how and when apologies work and why people apologize. In order to answer this question we conducted a laboratory experiment. We designed an experiment where people can harm others and where apologies were appropriate and reasonable. The great advantage of our design is that it controls for clearly intentional offenses and offenses that can be committed intentionally or due to inability.

Our results clearly show that the effect of an apology depends on the ambiguous intentionality of the offense. In order to make an apology work there has to be a positive probability that the offense has been committed unintentionally. Then the harmdoer can use the apology to convince the victim of his good intentions. If the offense was clearly intentional, an apology is useless. In this situation offenders do better not to apologize since an apology increases punishment. Apologies in this case seem to be interpreted as an affront. Offenders foresee this behavior and rarely apologize.

If the intentionality behind the offense is ambiguous we find that offenses are often followed by apologies – in particular if the harmdoer faces possible punishment. Harmdoers use apologies mainly in order to avoid punishment and not for reasons of remorse. Nevertheless, the harmdoers' strategy works: After offenses with ambiguous intentionality victims punish apologizers less than non-apologizers. Against conventional beliefs, our results show that the driving force behind this difference is not the apology's decreasing effect on punishment. It is rather the refusal to apologize that increases punishment and therefore causes the difference. A possible explanation is that victims expect an apology for the offense and punish if they do not receive one.

Our results show that the costs of an apology do not affect punishment behavior and therefore contradict predictions by Ohtsubo and Watanabe (2009) and Ho (2007). Costly apologies do not appear to be more credible than costless apologies.

With respect to the functioning of apologies we find that apologizing decreases the probability of being punished but does not reduce the extent of punishment if punishment occurs. People either accept an apology and therefore stop punishing, or they do not accept the apology and punish nevertheless. People totally forgive or do not forgive at all. Offenses are not partly unintentional and therefore you cannot "forgive just a little bit".

Appendix-Instructions

Instructions- Player A

Today you are participating in an economic experiment. By reading the following instructions carefully, you can – depending on your decisions – earn money in addition to the 2 euro show-up fee. Therefore it is important to read the instructions carefully. During the experiment it is not allowed to communicate with other participants. That is why we ask you not to talk with each other. If you have any questions, please take another look at the instructions. If you still have a question, please raise your hand. We will then come to you and answer your question in private. During the experiment we do not use euros, but points. All points you receive during the experiment will be changed into euros at the end of the experiment: 1 point =0.01 euros. The following pages give you instructions on the course of the experiment. At the end of the instructions you will find some control questions that will help you to understand the experiment. The experiment will start as soon as all participants are familiar with the experiment. **SUMMARY:** The experiment lasts 10 rounds. In every round you and a player B form a team. Both of you receive 2 questions. Answering these questions correctly changes your points. There are solo questions and team questions. Answering a solo question correctly gives 5 points. Answering a team question correctly means losing 40 points and gains 120 points for the team member. First player A solves team question A. Next, player B solves team question B. After team question B player B can send a message to you. Afterwards you can deduct points from player B. Then a new period starts. In every period you form a team with another player B. At the end of the experiment you receive a 2 euro show-up fee additionally to all points you receive during the experiment. **EXPERIMENT:** At the beginning of every round every player receives an endowment of 60 points. Next, every player A is matched with a random player B. You form a team for one round and answer questions. There are team questions and solo questions. **Team questions:** Every team receives 2 team questions: team question A and team question B. First, player A answers team question A, and then player B answers team question B. Answering a team question correctly means losing 40 points and gaining 120 points for the team member. This is clarified in the next table:

team question A	player A answers team question A correctly	player A answers team question A incorrectly
points for player A	-40	0
points for player B	120	0

It is exactly the opposite with team question B.

team question B	player B answers team question B correctly	player B answers team question B incorrectly
points for player A	120	0
points for player B	-40	0

Solo questions: When you are answering your team question, player B receives the same question as a solo question. By answering his solo question correctly, he receives 5 points. When player B is answering his team question, you receive the same question as a solo question. By answering your solo question correctly, you receive 5 points. After every question you learn whether the team question has been answered correctly and your corresponding points. When you had to solve a solo question, you also learn if you have answered the solo question correctly. Total points are calculated by the initial endowment of 60 points plus the points gained in team questions A and B. After team question B player B can send you a message. [Only in costly treatment: This message costs 5 points.] As soon as player B sent you the message, the message appears on your screen. If player B does not send a message, you will be informed too. [Only in punishment treatments: After receiving the message, you can deduct points from player B. You can deduct up to 100 points but not more points than player B owns. By deducting points form player B, these points are erased. Deducting 5 points cost 1 point, deducting 1 point costs 0.2 points]. Now the period is over. Every player A is matched with a new player B. After 10 rounds you will see a screen that shows your income from all periods.

Instructions- Player B

Today you are participating in an economic experiment. By reading the following instructions carefully, you can – depending on your decisions – earn money in addition to the 2 euro show-up fee. Therefore it is important to read the instructions carefully. During the experiment it is not allowed to communicate with other participants. That is why we ask you not to talk with each other. If you have any questions, please take another look at the instructions. If you still have a question, please raise your hand. We will then come to you and answer your question in private. During the experiment we do not use euros, but points. All points you receive during the experiment will be changed into euros at the end of the experiment: 1 point =0.01 euros. The following pages give you instructions on the course of the experiment. At the end of the instructions you find some control questions that will help you to understand the experiment. The experiment will start as soon as all participants are familiar with the experiment. **SUMMARY:** The experiment lasts 10 rounds. In every round you and a player A form a team. Both of you receive 2 questions. Answering these questions correctly changes your points. There are solo questions and team questions. Answering a solo question correctly means 5 points. Answering a team question correctly means losing 40 points and gaining 120 points for the team member. First player A solves team question A. Next, player B solves team question B. After team question B player B can send a message to you. Afterwards you can deduct points from player B. Then a new period starts. In every period you form a team with another player B. At the end of the experiment you receive a 2 euro show-up fee additionally to all points you receive during the experiment. **EXPERIMENT:** At the beginning of every round every player receives an endowment of 60 points. Next, every player A is matched with a random player B. You form a team for one round and answer questions. There are team questions and solo questions. **Team questions:** Every team receives 2 team questions: team question A and team question B. First, player A answers team question A, and then player B answers team question B. Answering a team question correctly means losing 40 points and gaining 120 points for the team member. This is clarified in the next table:

team question A	player A answers team question A correctly	player A answers team question A incorrectly
points for player A	-40	0
points for player B	120	0

It is exactly the opposite with team question B.

team question B	player B answers team question Bcorrectly	player B answers team question Bincorrectly
points for player A	120	0
points for player B	-40	0

Solo questions: When you are answering your team question, player B receives the same question as a solo question. By answering his solo question correctly, he receives 5 points. When player B is answering his team question, you receive the same question as a solo question. By answering your solo question correctly, you receive 5 points. After every question you learn if the team question has been answered correctly and your corresponding points. When you had to solve a solo question, you also learn if you have answered the solo question correctly. After team question B you can send a message to player A. [Only in costly treatment: This message costs 5 points.] [Only in punishment treatments: After receiving the message, player A can deduct points from you. He can deduct up to 100 points but not more points than you own. Deducting points from you means that these points are erased. Deducting 5 points cost 1 point, deducting 1 point costs 0.2 points.] Now the period is over. Every player A is matched with a new player B. After 10 rounds you will see a screen that shows your income from all periods.

7. Literature

ABELER, J., CALAKI, J., ANDREE, K. and BASEK, C., 2009. The Power of Apology, *CeDEx Discussion Paper Series*,

BENNETT, M. and EARWAKER, D., 1994. Victims' Responses to Apologies: The Effects of Offender Responsibility and Offense Severity, *Journal of Social Psychology*, 134 (4), pp.457-464.

BLOUNT, S., 1995. When social outcomes aren't fair: The effect of causal attributions on preferences., *Organizational Behavior and Human Decision Processes*, 63 (2), pp.131 - 144.

BOLTON, G. E. and OCKENFELS, A., 2000. ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review*, 90 (1), pp.166-93.

BOTTOM, W. P., GIBSON, K., DANIELS, S. E. and MURNIGHAN, J. K., 2002. When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation, *Organization Science*, 13 (5), pp.497-513.

BRANDTS, J. and SOLA, C., 2001. Reference Points and Negative Reciprocity in Simple Sequential Games, *Games and Economic Behavior*, 36 (2), pp.138-57.

CHARNESS, G. and DUFWENBERG, M., 2006. Promises and partnership, *Econometrica*, 74 (6), pp.1579-1601.

CHARNESS, G. and LEVINE, D. I., 2007. Intention and stochastic outcomes: An experimental study, *Economic Journal*, 117 (522), pp.1051-1072.

COHEN, J. R., 2002. Legislating apology, *University of Cincinatti Law Review*, 70 819-872.

ELLINGSEN, T. and JOHANNESSON, M., 2008. Anticipated verbal feedback induces altruistic behavior, *Evolution and Human Behavior*, 29 (2), pp.100-105.

EXLINE, J. J., DESHEA, L. and HOLEMAN, V. T., 2007. Is apology worth the risk? Predictors, outcomes, and ways to avoid regret, *Journal of Social & Clinical Psychology*, 26 (4), pp.479-504.

FALK, A., FEHR, E. and FISCHBACHER, U., 2003. On the nature of fair behavior, *Economic Inquiry*, 41 (1), pp.20-26.

FALK, A., FEHR, E. and FISCHBACHER, U., 2008. Testing Theories of Fairness-Intentions Matter, *Games and Economic Behavior*, 62 (1), pp.287 - 303.

FEHR, E. and SCHMIDT, K. M., 1999. A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114 (3), pp.817-68.

FEIN, S., 1996. Effects of suspicion on attributional thinking and the correspondence bias, *Journal of Personality and Social Psychology*, 70 (6), pp.1164-1184.

FISCHBACHER, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics*, 10 (2), pp.171-178.

FRANTZ, C. and BENNIGSON, C., 2005. Better late than early: The influence of timing on apology effectiveness, *Journal of Experimental Social Psychology*, 41 (2), pp.201-207.

GIRARD, M., MULLET, Ã. T. and CALLAHAN, S., 2002. Mathematics of forgiveness, *American Journal of Psychology*, 115 (3), pp.351-375.

GNEEZY, U. and RUSTICHINI, A., 2000. Pay enough or don't pay at all, *Quarterly Journal of Economics*, 115 (3), pp.791-810.

GOFFMAN, E., 1971. *Relations in public*, Harper Colophon Books, New York.

GOLD, G. and WEINER, B., 2000. Remorse, confession, group identity, and expectancies about repeating a transgression, *Basic and Applied Social Psychology*, 22 (4), pp.291-300.

GREINER, B., 2004. The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. University of Cologne, Department of Economics.

HO, B., 2007. Apologies as signals, *Cornell University Discussion Paper*,
HURKENS, S. and KARTIK, N., 2009. Would I lie to you? On social preferences and lying aversion, *Experimental Economics*, 12 (2), pp.180-192.
LEVINE, D. K., 1998. Modeling Altruism and Spitefulness in Experiments, *Review of Economic Dynamics*, 1 (3), pp.593-622.
LUNDQUIST, T., ELLINGSEN, T., GRIBBE, E. and JOHANNESSON, M., 2009. The aversion to lying, *Journal of Economic Behavior and Organization*, 70 (1), pp.81-92
MCCULLOUGH, M. E., RACHAL, K. C., SANDAGE, S. J., WORTHINGTON, E. L., JR., BROWN, S. W. and HIGHT, T. L., 1998. Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement, *Journal of Personality and Social Psychology*, 75 (6), pp.1586-1603.
MCCULLOUGH, M. E., WORTHINGTON, E. L., JR. and RACHAL, K. C., 1997. Interpersonal forgiving in close relationships, *Journal of Personality and Social Psychology*, 73 (2), pp.321-336.
OHBUCHI, K.-I., KAMEDA, M. and AGARIE, N., 1989. Apology as aggression control: Its role in mediating appraisal of and response to harm, *Journal of Personality and Social Psychology*, 56 (2), pp.219-227.
OHBUCHI, K.-I. and SATO, K., 1994. Children's Reactions to Mitigating Accounts: Apologies, Excuses, and Intentionality of Harm, *Journal of Social Psychology*, 134 (1), pp.5-17.
OHTSUBO, Y. and WATANABE, E., 2009. Do sincere apologies need to be costly? Test of a costly signaling model of apology, *Evolution and Human Behavior*, 30 (2), pp.114-123.
SÁNCHEZ-PAGÉS, S. and VORSATZ, M., 2007. An experimental study of truth-telling in a sender-receiver game, *Games and Economic Behavior*, 61 (1), pp.86-112.
SÁNCHEZ-PAGÉS, S. and VORSATZ, M., 2009. Enjoy the silence: An experiment on truth-telling, *Experimental Economics*, 12 (2), pp.220-241.
SCHER, S. J. and DARLEY, J. M., 1997. How Effective Are the Things People Say to Apologize? Effects of the Realization of the Apology Speech Act, *Journal of Psycholinguistic Research*, 26 (1), pp.127-140.
SCHLENKER, B. R. and DARBY, B. W., 1981. The use of apologies in social predicaments, *Social Psychology Quarterly*, 44 (3), pp.271-278.
SCHMITT, M., GOLLWITZER, M., FÖRSTER, N. and MONTADA, L., 2004. Effects of objective and subjective account components on forgiving, *The Journal of social psychology*, 144 (5), pp.465-486.
SCHUL, Y., MAYO, R. and BURNSTEIN, E., 2004. Encoding under trust and distrust: The spontaneous activation of incongruent cognitions, *Journal of Personality and Social Psychology*, 86 (5), pp.668-679.
SCHWEITZER, M. E., HERSHEY, J. C. and BRADLOW, E. T., 2006. Promises and lies: Restoring violated trust, *Organizational Behavior and Human Decision Processes*, 101 (1), pp.1-19.
SHAPIRO, D., 1991. The effects of explanations on negative reactions to deceit, *Administrative Science Quarterly*, 36 (4), pp.614-630.
SHAW, J. C., WILD, E. and COLQUITT, J. A., 2003. To justify or excuse?: A meta-analytic review of the effects of explanations, *Journal of Applied Psychology*, 88 (3), pp.444-458.
SKARLICKI, D. P., FOLGER, R. and GEE, J., 2004. When Social Accounts Backfire: The Exacerbating Effects of a Polite Message or an Apology on Reactions to an Unfair Outcome, *Journal of Applied Social Psychology*, 34 (2), pp.322-341.
STRUTHERS, C. W., EATON, J., SANTELLI, A. G., UCHIYAMA, M. and SHIRVANI, N., 2008. The effects of attributions of intent and apology on forgiveness: When saying

sorry may not help the story, *Journal of Experimental Social Psychology*, 44 (4), pp.983-992.

TAKAKU, S., WEINER, B. and OHBUCHI, K.-I., 2001. A cross-cultural examination of the effects of apology and perspective taking on forgiveness, *Journal of Language and Social Psychology*, 20 (1), pp.144-166.

TAVUCHIS, N., 1991. *Mea culpa: A sociology of apology and reconciliation*, Mea culpa: A sociology of apology and reconciliation.: Stanford University Press.

UTIKAL, V. and FISCHBACHER, U., 2009. On the attribution of externalities, *TWI Research Paper Series*,

WADA, M., 1998. The Mitigative Effects of an Apology and Excuses in a Social Predicament, *Journal of Yamanashi Eiwa Jinior College*, 32 ( ), pp.84-72.

WEINER, B., GRAHAM, S., PETER, O. and ZMUIDINAS, M., 1991. Public confession and forgiveness, *Journal of Personality*, 59 (2), pp.281-312.

WHITE, B. T., 2009. Saving face, *Arizona Legal Studies Discussion paper*,

WITVLIET, C., WORTHINGTON JR, E. and WADE, N., 2002. Victims' heart rate and facial EMG responses to receiving an apology and restitution, *Psychophysiology*, Supplement (2002), pp.88.

ZECHMEISTER, J. and ROMERO, C., 2002. Victim and offender accounts of interpersonal conflict: Autobiographical narratives of forgiveness and unforgiveness, *Journal of Personality and Social Psychology*, 82 (4), pp.675-686.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch