Communication and Guilt in a Prisoner's Dilemma

Miettinen, T.; Suetens, S.

Publication date:
2008

CentER

Discussion Paper

No. 2008–12

COMMUNICATION AND GUILT IN A PRISONER'S DILEMMA

By Topi Miettinen, Sigrid Suetens

February 2008

ISSN 0924-7815

Tilburg University

Communication and guilt in a prisoner's dilemma

Topi Miettinen*and Sigrid Suetens†

February 5, 2008

Abstract

In this paper we measure experienced guilt in a prisoner's dilemma experiment with pre-play communication. We find that feelings of guilt only arise in the case of unilateral defection and that they are stronger when players have mutually agreed to cooperate. We also find that fining unilateral defection reduces feelings of guilt.

Keywords: cooperation, guilt, experiments, prisoner's dilemma, pre-play communication

JEL codes: C91, D74

*Max Planck Institute of Economics, Kahlaische Straße 10, 07745 Jena, Germany.

†CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. The authors thank Raúl López-Pérez, Jan Potters and Ondrej Rydval for useful suggestions.

1 Introduction

A robust finding in social dilemma experiments is that communication prior to play of the game increases cooperation dramatically (see Ledyard, 1995; Sally, 1995; Ostrom, 2006; Bicchieri and Lev-On, 2007, for overviews). Social psychologists, behavioral economists and sociologists alike have striven to understand the fundamental motivation behind this effect.

Researchers have reached a rather undisputed consensus about the prime driving force of the beneficial effect of communication on cooperation. Communication is seen as something enabling parties to establish an informal agreement on which actions should be taken.[1] Orbell et al. (1988) and Kerr and Kaufman-Gilliland (1994), for example, show that successful communication often consists of making promises and engaging in a commitment to cooperate. Bicchieri and Lev-On (2007) conclude, on the basis of an overview of experiments on communication and cooperation, that communication makes people focus on pro-social norms and allows them to develop mutual beliefs about behavior.

This paper contributes to this literature by measuring how experienced guilt reacts to agreements and choices in such a setting. Subjects in our experiment play an anonymous, one-shot prisoner's dilemma game with monetary incentives. Prior to play, they have an opportunity to communicate a willingness to cooperate. After the actions in the dilemma have been chosen, guilt valences are elicited twice, both before and after revealing the matched participant's choice.

Guilt is viewed as a mediation emotion by social psychologists. Baumeister et al. (1994), in particular, adopt the view that the main function of guilt is interpersonal: it enforces communal norms by providing incentives to avoid transgressions that harm others. Based on this view, the measurement of guilt may help to identify whether and how strong a norm a certain behavioral profile is. In a prisoner's dilemma experiment, it is natural to focus on the norm of cooperation and the impact of a mutual cooperative agreement on this cooperative norm.

The importance of guilt as such a mediation and enforcement emotion has been highlighted in recent studies in social psychology and in economics.[2] In

[1] Other potential effects of communication such as, for example, decreasing social distance, increasing solidarity and creation of group identity have been shown to be at most partial explanations (see, e.g., Kerr and Kaufman-Gilliland, 1994; Frohlich and Oppenheimer, 1998; Brosig et al., 2003; Bochet et al., 2006). Bicchieri (2002) views creating group identity in itself as a trigger for norm-abiding behavior.

[2] See Eisenberg (2000) for an overview of social psychology studies and see, for instance, Akerlof (1980); Dufwenberg and Gneezy (2000); Ellingsen and Johanneson (2004);

their guilt-aversion model, Charness and Dufwenberg (2006) assume that promises convey information about intended choices and that agents do not want to let down the other by acting counter to what the other expects them to do. Miettinen (2006) and Lopez-Perez (2008) assume that transgressions of a cooperative norm only induce feelings of guilt *when others stick to the norm*. Experimental studies have shown that guilt and shame indeed play a crucial role in social situations. Ketelaar and Au (2003) show that people who feel guilty after having defected in a social dilemma game tend to cooperate more in further repetitions of the game. Hopfensitz and Reuben (2007) find that someone who is punished for having defected, will not retaliate in the future only if she feels guilty or ashamed.

We find that people only feel guilty about defecting if the partner has not defected as well. We also find that guilt feelings about defecting are stronger when a mutual agreement to cooperate has been reached than without such agreement. Moreover, when a small monetary fine is imposed on unilateral defectors, guilt feelings of unilateral defectors are significantly reduced. This latter pattern is consistent with the result from previous experiments (see, e.g., Eisenberger and Cameron, 1996; Gneezy and Rustichini, 2000a,b; Fehr and Rockenbach, 2003; Falk and Kosfeld, 2006), that an extrinsic incentive to adhere to norm-abiding behavior might crowd-out the intrinsic incentive to stick to a pro-social norm.

The paper is organized as follows. Section 2 presents the experimental design and procedure and the research hypotheses. Section 3 has the results. Section 4 concludes.

2 Design, procedure and hypotheses

The experiment was run in January 2007 in the computer laboratory of the Max Planck Institute of Economics with a total of 140 students from different fields of study at the Friedrich-Schiller University of Jena, Germany. Participants were recruited using the ORSEE software (Greiner, 2004) and the experiment was programmed with *z-Tree* software (Fischbacher, 2007).

After being seated at a visually isolated computer terminal, participants received written instructions (see appendix A for an English translation). Understanding of the rules was assured by a control questionnaire that subjects had to answer before the experiment could start.

Each participant was randomly paired with another participant and played a simple prisoner's dilemma (PD) game with payoffs in Euro as depicted in

Charness and Dufwenberg (2006); Miettinen (2006); Lopez-Perez (2008) in the theoretical economics literature.

table 1. Eventually, they were asked to choose option 1 (defect) or option 2 (cooperate). Yet, before entering the choices in the PD, each participant had the choice between either not communicating or communicating to the other that she would like them both to cooperate.[3] They were told that communication was not binding. After the communication phase, subjects received information about the matched participant's willingness to engage in a cooperative agreement.

When choices in the PD had been made, a new computer screen appeared where first- and second-order beliefs were elicited. We preferred simplicity over precision and, thus, each participant was paid an additional 0.5 Euro, first, for correctly predicting the choice of the participant they were paired with, and second, for correctly predicting the other's prediction of one's own choice. Total earnings were revealed at the end of the experiment only (the sum of earnings from the PD game and earnings from predicting the matched participant's choice and prediction).

Beliefs were elicited to allow us to deal with a potential reverse causality between feelings of guilt and decisions in the PD. If guilt enforces agreements, then it is the incentive to avoid guilt feelings that keeps agents from breaching. Thus rather than the outcome causing guilt valences, the expected guilt valences might cause the outcome. To study the robustness of our results, we circumvent the issue of reverse causality by studying players who expect their opponent to defect. These players should not expect to suffer guilt (see end of Section 3.2). Moreover, eliciting beliefs is also interesting because it allows us to shed light on whether and how players expect communication to matter (see Section 3.1).

Emotional reactions of guilt were first measured immediately after the choices and belief elicitation and *before* revealing the matched participant's choice. Emotional reactions were measured another time immediately *after* revealing the matched participant's choice. In both measurements, subjects filled in their emotional valence on a seven-point scale starting from "not at all" to "very intensely" (see also Ketelaar and Au, 2003; Hopfensitz and Reuben, 2007).[4] Self-reports are generally considered to be a valid technique to measure (Robinson and Clore, 2002). This is the case for guilt in particular since guilt is difficult to measure using physiological or neurological measurement (Adolphs, 2002).

Measurements before and after allow us to study within subjects how the

[3]They were asked to choose (simultaneously) between the following two options: "I communicate to the other that I want us both to choose option 2" and "I prefer not to communicate".

[4]We also measured other emotions such as shame, sadness, happiness, anger, etc., in order to avoid revealing the aim of the study.

	your payoff	the other's payoff
you choose option 1 and the other chooses option 1	6	6
you choose option 1 and the other chooses option 2	16	4
you choose option 2 and the other chooses option 1	4	16
you choose option 2 and the other chooses option 2	10	10

Table 1: Payoffs in the prisoner's dilemma game

feeling of guilt due to defection is influenced by learning that the matched participant defected as well. It is often assumed that guilt due to defection only arises when the other does not defect as well (see Bicchieri and Erte, 2007; Lopez-Perez, 2008; Miettinen, 2006). Our design allows us to directly verify this assumption. Our first hypothesis is the following.

Hypothesis 1 *Guilt feelings mainly arise after unilateral defection.*

According to a norm-based hypothesis, communication resulting in mutual agreements increases cooperation because it induces or enforces the norm of cooperation. Deviating from the norm gives rise to guilt feelings. This is our second hypothesis.

Hypothesis 2 *Guilt feelings due to (unilateral) defection are stronger when there is a mutual agreement to cooperate than when there is no agreement.*

Finally, we also study whether guilt after unilateral defection is reduced when an extrinsic incentive to mutually cooperate is provided (see, e.g., Gneezy and Rustichini, 2000a). To study this question, we introduced a penalty for unilateral defectors in part of the treatments. In the first of these treatments, the instructions included a sentence stating that, when one chooses option 1 (i.e., defects) while the other chooses option 2 (i.e., cooperates), an amount of 0.20 Euro would be transferred to the other player. In the second of these treatments, the instructions were the same, but there was a computer screen that popped up after the communication phase to remind the participants of the possible penalty. In the third treatment the penalty was 1 Euro. Table 2 gives an overview of all treatments including the number of participants. According to the crowding-out hypothesis, a small penalty reduces feelings of guilt due to unilateral defection.

Hypothesis 3 *Guilt feelings due to unilateral defection are weaker when a small penalty for unilateral defection is introduced than when it is not introduced.*

treatment 1	no penalty	44
treatment 2	penalty 0.20	32
treatment 3	penalty 0.20 and reminder screen	32
treatment 4	penalty 1 and reminder screen	32

Table 2: Treatments and number of participants

3 Results

3.1 Cooperation, communication and beliefs

In this subsection, we give a descriptive overview of average behavior, communication and beliefs. In subsection 3.2. we subject our research hypotheses to tests and discuss our findings.

Figure 1 gives an overview of cooperation rates. The rates are conditioned on whether communication has resulted in a mutual agreement, and on whether beliefs are cooperative or not (as defined further below). We say that a mutual agreement to cooperate is established when both parties made use of the communication device to signal a willingness to cooperate.

Figure 1 shows that the overall cooperation rate is 50%. The cooperation rate in cases with a mutual agreement is 59%, which is significantly higher than the 26% without agreement ($p = .001$ in χ^2-test).[5] Still, even with mutual agreements a significant number of subjects defected, which allows us to study the impact of agreements on guilt.

As figure 1 shows, beliefs are more cooperative if a mutual agreement has been reached. We say that one has *cooperative beliefs* if one believes that the other has cooperated and that the other believes one has cooperated (i.e., if both first- and second-order beliefs are cooperative). It is clearly visible that subjects do not ignore beliefs as they should if they followed the monetary incentive to play the dominant strategy to defect, and they do not believe that others ignore beliefs either. Out of the 102 subjects with a mutual agreement, 72 have cooperative beliefs, compared to only 9 out of the 38 subjects without a mutual agreement (significant with $p = .000$ in χ^2-test). A total of 74% out of the subjects with a mutual agreement *and* cooperative beliefs actually cooperate, which is higher than the 23% cooperation rate of subjects with a mutual agreement but different beliefs (significant with

[5]We do not claim that the significant differences in this section on descriptives are direct tests of the effect of communication on cooperation or beliefs, nor a test of communication or beliefs on behavior. Agents may self-select into the no-agreement condition, for instance. For direct tests, we refer to existing literature on communication and cooperation overviewed by, for instance, Sally (1995) and Bicchieri and Lev-On (2007).

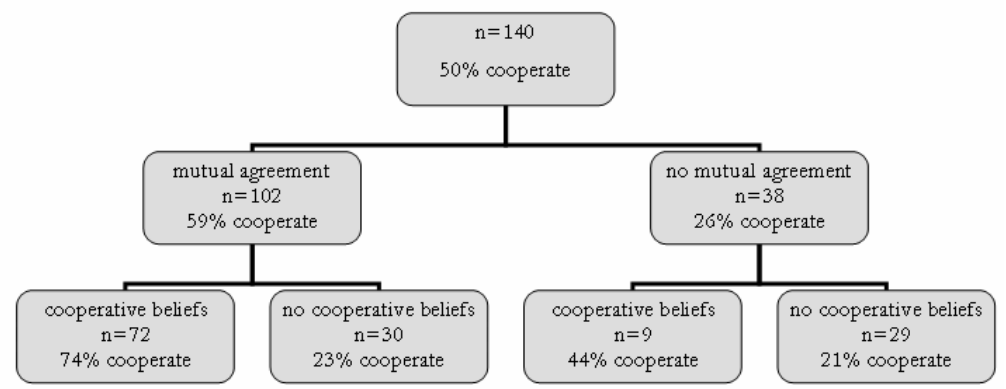


Figure 1: Overview of cooperation rates

$p = .000$ in χ^2-test), and higher than the 44% cooperation rate of subjects without a mutual agreement and with cooperative beliefs (significant with $p = .000$ in χ^2-test). Cooperation rates of subjects without cooperative beliefs and either with or without a mutual agreement are not significantly different (23% versus 21%).

In order to understand how communication relates to first- and second-order beliefs, we provide a more detailed picture on communication and beliefs. Table 3 reports results from two probit regressions where beliefs are regressed on a dummy variable indicating whether one has communicated or not (self communicate) and one indicating whether the matched subject has communicated or not (other communicate). On the one hand, communication of the other is significantly positively correlated with the belief that the other will cooperate (it increases by 55%) and is not correlated with the second-order belief. One's own communication, on the other hand, is significantly correlated with the second-order belief (increases by 26%) and not with the first-order belief. Thus, the first-level effects of communication are larger and significant: the other's communication makes one think that the other is more likely to cooperate and gives a conditional cooperator a stronger incentive to cooperate.[6] Similarly, one's own communication is mainly related to one's second-order belief and not to one's first-order belief.

Table 4 gives a general overview of cooperation and communication rates in all treatments. The communication rate measures the percentage of subjects that chose to communicate. Cooperation rates are not significantly different between the treatments with and without penalty ($p = .585$ in χ^2-

[6]On conditional cooperation, see, for instance, Keser and van Winden (2000) and Fischbacher et al. (2002).

left-hand-side	self communicate est (p-value)	other communicate est (p-value)	log l	pseudo R^2
1st-order beliefs	0.17 (.153)	0.55 (.000)***	-77.52	.13
2nd-order beliefs	0.26 (.054)*	0.21 (.128)	-85.48	.04

***, **, * Significant at the 1%, 5% and 10% level (two-tailed tests).
Standard errors are robust for potential within-pair dependency.
Number of observations is 140.
Marginal effects are reported.

Table 3: Effects of communication on beliefs

treatment	penalty	cooperation rate	communication rate	# obs
1	no	45% (20)	93% (41)	44
2	yes	53% (17)	78% (25)	32
3	yes	56% (18)	84% (27)	32
4	yes	47% (15)	84% (27)	32
avg 2,3,4	yes	52% (50)	82% (79)	96
avg all	-	50% (70)	86% (120)	140

Table 4: Cooperation and communication rates across treatments

test), nor are communication rates ($p = .119$ in χ^2-test). The introduction of a small penalty has thus not reduced cooperation. Therefore in our experiment, unlike in Gneezy and Rustichini (2000b), for instance, there is no crowding-out effect on *behavior*. In the next subsection, we find that there is a crowding-out effect on *the emotional valence of guilt*, however.

3.2 Guilt, norms and crowding-out

In this section we examine our research hypotheses. Table 5 gives a summary of median values of guilt intensities measured before and after receiving feedback about the other's choice.[7] We distinguish between cases in which one defects or cooperates and cases in which one has a mutual agreement or not. Median guilt intensities measured *after* feedback are additionally classified based on whether the matched partner defects or cooperates. A guilt intensity of 0 refers to not feeling guilty at all and an intensity of 6 refers to feeling it very intensely.

[7]Since guilt intensities are measured on an ordinal scale, the median is the appropriate measure for central tendency. See appendices B and C for frequency distributions of guilt intensities measured before and after feedback.

	before feedback	**after feedback**	
		other defects	other cooperates
participant defects			
no mutual agreement	0	0	1
	n=28	n=20	n=8
mutual agreement	2	0	2
	n=42	n=20	n=22
participant cooperates			
no mutual agreement	0	0	0
	n=10	n=8	n=2
mutual agreement	0	0	0
	n=60	n=22	n=38

Table 5: Median guilt intensities

The table clearly shows that feelings of guilt exclusively arise when a subject defects. More specifically, according to the second measurement of guilt valence when the choice of the matched partner is revealed, the guilt feelings arise only when the subject defects and the partner cooperates. This is clearly in line with hypothesis 1. A within-subjects Wilcoxon signed-ranks test confirms that feelings of guilt after defection are significantly less strong once participants learn that the counterpart also defected ($p = .000$). The table further also shows that median intensities of guilt about unilaterally defecting (measured before and after feedback) are higher with a mutual agreement to cooperate than without such an agreement. This is in line with hypothesis 2.

In order to statistically test research hypotheses 1, 2 and 3 we perform six univariate ordered probit regressions where intensities of guilt measured before and after feedback are regressed on a single binary variable of interest. This technique has the advantage that standard errors of the estimated regression coefficients can be estimated such that they are robust to within-pair dependency without losing the individual observations.[8] This technique is very useful for analyzing our data since — given the pre-play communication phase — choices within pairs cannot be assumed to be independent. Mann-Whitney U-tests, for instance, do not allow taking into account within-pair dependency without losing individual observations.

Table 6 contains the six estimated coefficients and associated p-values in corresponding six univariate regressions where guilt intensity is the left-

[8]Robust estimation of standard errors is done using the so-called sandwich estimator due to Huber (1967) and White (1982).

hand-side variable and 'mutual agreement', 'other cooperates' or 'penalty' is the right-hand-side variable. The variables 'mutual agreement' and 'other cooperates' are self-explanatory. The variable 'penalty' indicated whether the treatment in question was one with an exogenous penalty on unilateral deviations. This variable is included in order to test hypothesis 3. Note that a defecting subject is only punished in treatments 2 to 4 if she unilaterally defects.[9] Column (1) studies all cases where the subject chose to defect. Column (2) studies only those defectors whose opponent chose to cooperate. Column (3) studies those of the latter who moreover expected the other to cooperate. Thus, the defectors in the third column have counterfactual expectations.

The estimated coefficients do not represent marginal effects, but indicate how the intensity of guilt (measured either before or after feedback) is related to the right-hand-side variable of interest. A positive and significant coefficient of, say, 'mutual agreement' implies that the probability of observing a high (low) guilt intensity is higher (lower) with than without a mutual agreement to cooperate.

First, we focus on the two leftmost columns of Table 6, i.e. columns (1) and (2). The rightmost column (3) is analyzed further below. The second coefficient in the column (1) shows that a defector's probability of feeling more guilty is significantly higher when the other has cooperated than when the other has defected as well ($p < .001$). This further supports hypothesis 1.

Second, concerning hypothesis 2, the table illustrates two features. On the one hand, the first coefficient in column (1) shows that a subject with a mutual agreement to cooperate is significantly more likely to have strong guilt feelings about defecting ($p < .010$, based on before-feedback measurement). On the other hand, the first coefficient in column (2) shows that this is true also for *unilaterally* defection ($p < .050$, based on after-feedback measurement). Hypothesis 2 is thus supported as well.

Finally, since the second coefficient in column (2) is negative and significant, also hypothesis 3 is supported by the data. Thus, a penalty for unilateral defection reduces the probability of having strong guilt feelings about it ($p < .050$).[10]

[9]We do not regress guilt measured *before* feedback on a penalty variable since at the point of measurement a deviating subject does not know yet whether she will be actually punished. The coefficient in a regression of guilt before feedback on a 'potential penalty' variable, which measures whether a penalty is potentially introduced (i.e., when one defects in treatments 2 to 4), is not significant ($p = .356$).

[10]Regressions using shame instead of guilt as a dependent variable produce similar results with respect to the effect of unilateral defection and the effect of a penalty (cf.

	(1)	(2)	(3)
	n=70	n=30	n=15
guilt before feedback			
mutual agreement	0.78 (.006)***	-	
guilt after feedback			
other cooperates	1.38 (.000)***	-	
mutual agreement	-	0.74 (.045)**	1.16 (.042)**
penalty	-	-0.93 (.017)**	-0.71 (.116)

(1) Given defection.
(2) Given defection and other's cooperation.
(3) Given defection and other's cooperation by surprise.
*** and ** Significant at the 1% and 5% level (one-tailed tests).
p-values are in brackets.

Table 6: Estimated coefficients of six regressions of guilt (including p-values)

Let us now turn to column (3) in table 6 which tackles the issue of reverse causality. Although we are primarily interested in experienced guilt, which is measured after players have made choices in the PD, choices may very well be motivated by avoidance of expected guilt (see, e.g., Dufwenberg, 2002; Bicchieri and Erte, 2007; Miettinen, 2006; Lopez-Perez, 2008). That anticipated guilt influences choices may weaken causal effects such that coefficients may reflect mere correlations. Yet, our design allows us to circumvent this problem.

As table 5 shows, there is no guilt if the other defects. This feature of guilt, together with the fact that we elicit beliefs after choices and before revealing outcomes, allows us to tackle this endogeneity issue. There are 15 players who expected the other to defect, but surprisingly ended up with a cooperating partner and feeling guilty about defecting. We can safely assume that these players were not motivated by avoiding expected feelings of guilt when deciding to defect. That is, we can assume that in the latter case, the unexpected combination of own defection and the other's cooperation *causes* guilt feelings. The estimates presented in column (3) in table 6 are based on ordered probit regressions where only subjects who unilaterally defect and *expected the other to defect* are included. Based on these estimates, the 'mutual agreement' variable is still positive and significant ($p < .050$). The

hypotheses 1 and 3), but not with respect to the effect of a mutual agreement (cf. hypothesis 2). This is not surprising given that shame is taken to be a different emotion than guilt in recent studies in social psychology (see, e.g., Eisenberg, 2000; Smith et al., 2002; Fontaine et al., 2006; Tangney et al., 2007). Guilt is generally viewed as the 'more moral' emotion of the two.

'penalty' variable fails to be significant, however.

4 Conclusion

By studying reported guilt valences in a prisoner's dilemma with communication, we shed new light on why communication increases cooperation in those settings. We design a simple two-player prisoner's dilemma experiment where parties can communicate a willingness to cooperate prior to choosing whether to cooperate or to defect and where guilt valences are elicited before and after getting feedback about the matched person's action. In line with earlier experiments, our data show that the cooperation rate is higher when parties engage in a mutual agreement to cooperate.

Our focus is on how experienced guilt is related to choices and communication in the prisoner's dilemma. By measuring experienced guilt we aim to identify whether subjects interpret an agreement to cooperate as establishing and strengthening a normative code of cooperation. We view guilt as a mediation emotion that helps to understand which behavior is considered normatively ideal.

We find, first, that subjects experience guilt mainly when they defect, and more specifically when they *unilaterally* defect. Second, we find that guilt feelings due to (unilateral) defection are stronger with than without a mutual cooperative agreement. That is, we provide evidence in favor of mutual cooperation being considered to be a behavioral norm, the more so when a mutual agreement to cooperate is in place. Given that subjects learn to avoid choices that cause guilt (like in Ketelaar and Au, 2003; Hopfensitz and Reuben, 2007) and that the payoff from mutual cooperation is higher than the payoff from mutual defection[11], one should thus expect that communication increases cooperation rates. Finally, we also find that fining unilateral defection reduces feelings of guilt, which is in line with the so-called crowding-out findings in the literature.

We are not aware of previous experimental studies that survey emotions during or after the play of a social dilemma with communication. More generally, we see such emotional valence elicitation as a recommendable method to complement behavioral evidence to better understand the reasons behind the choice patterns in experiments.

[11] Another way of avoiding feelings of guilt is to refrain from committing to a cooperative agreement.

Appendix A: Translation of written instructions

Welcome to the experiment! The experiment is part of a research program financed by a research foundation. You will be asked to make a number of decisions. It is important that you follow the instructions carefully, because you can earn a considerable amount of money. At the end of the experiment, you will be paid your earnings in private and in cash. During the experiment you are not allowed to talk to other participants. If something is not clear, please raise your hand and one of us will help you.

In the experiment you will be matched with a randomly chosen participant. You both get the same instructions. What you will earn, depends on your own choice and the choice of this other participant. His/her identity will not be revealed to you, nor will your identity be revealed to him/her. The other participant will be referred to by "the other" in what follows. We ask you and the other to make a choice between option 1 and option 2. What you and the other will earn in Euro will be calculated as follows: see table 1.

Before entering your choices you and the other have the possibility to communicate to each other that you would like both to choose option 2 (computer screen called COMMUNICATION). We ask you to decide on this screen whether you prefer not to communicate or whether you communicate that you want both (you and the other) to choose option 2. After you sent your communication, you will be informed about what the other has communicated.

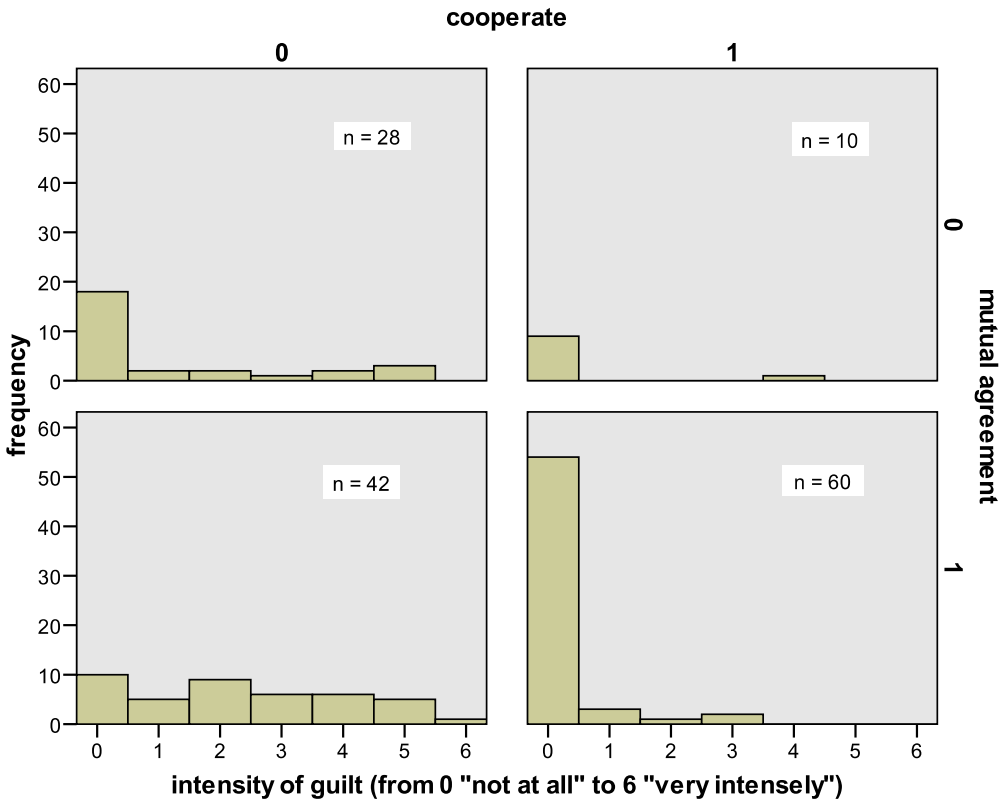
The communications are not binding.

[If you end up choosing option 1 and the other ends up choosing option 2, a penalty of 0,20 (1) Euro will be subtracted from your earnings and transferred to the other.] [only included in treatments 2, 3 and 4]

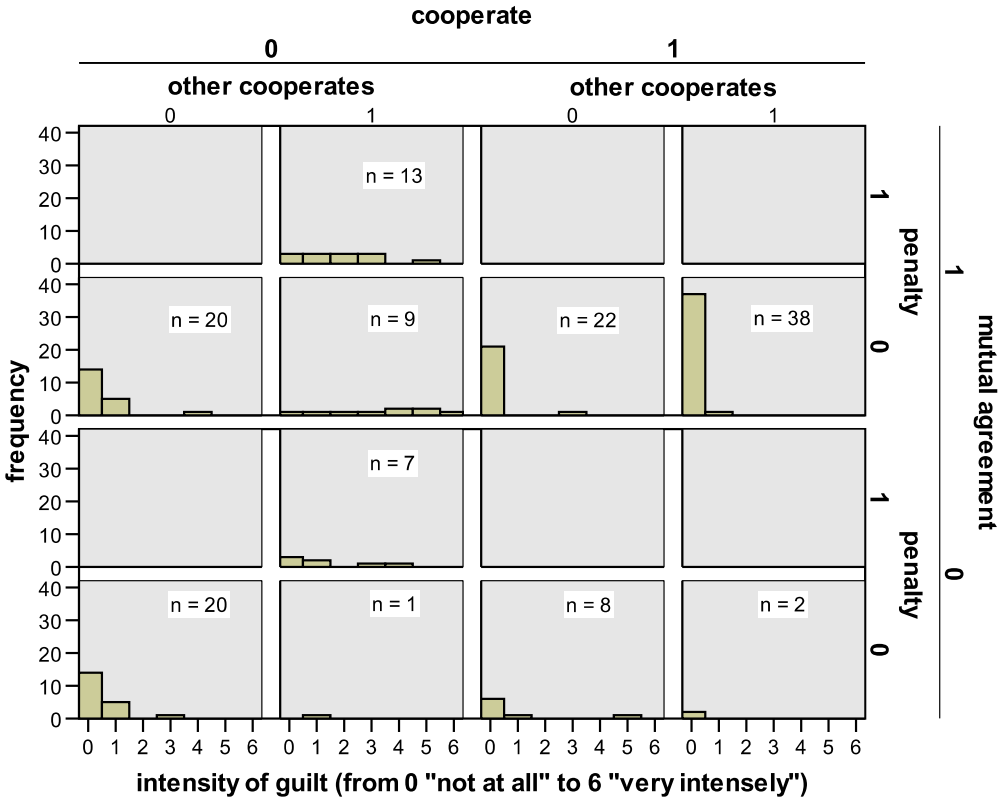
On the screen called CHOICE ENTRY you enter your final choice (option 1 or option 2).

Before the actual experiment starts, you will be asked to answer some control questions on the computer. This is to test whether you understand the payoff table. Please follow the instructions on the screen from now on.

Appendix B: Frequency distribution of guilt intensity measured *before* feedback



Appendix C: Frequency distribution of guilt intensity *after* feedback



Bibliography

R. Adolphs. Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1:21–61, 2002.

G. Akerlof. A theory of social customs, of which unemployment may be one consequence. *Quarterly Journal of Economics*, 94:749–775, 1980.

R. Baumeister, A. Stillwell, and T. Heatherton. Guilt: an interpersonal approach. *Psychological Bulletin*, 115:243–267, 1994.

C. Bicchieri. Covenants without swords: group identity, norms, and communication in social dilemmas. *Rationality & Society*, 14:192–228, 2002.

C. Bicchieri and X. Erte. *Do the right thing, but only if others do so*. 2007.

C. Bicchieri and A. Lev-On. Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. *Politics, Philosophy & Economics*, 6:139–68, 2007.

O. Bochet, T. Page, and L. Putterman. Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60:11–26, 2006.

J. Brosig, A. Ockenfels, and J. Weimann. The effect of communication media on cooperation. *German Economic Review*, 4:217–41, 2003.

G. Charness and M. Dufwenberg. Promises and partnerships. *Econometrica*, 74:1579–1601, 2006.

M. Dufwenberg. Marital contracts. *Journal of Economic Behavior & Organization*, 48:57–69, 2002.

M. Dufwenberg and U. Gneezy. Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, 30:163–83, 2000.

N. Eisenberg. Emotion, regulation, and moral development. *Annual Review of Psychology*, 51:665–97, 2000.

R. Eisenberger and J. Cameron. Detrimental effects of reward. Reality or myth? *American Psychologist*, 51:1153–1166, 1996.

T. Ellingsen and M. Johanneson. Promises, threats, and fairness. *Economic Journal*, 114:397–420, 2004.

A. Falk and M. Kosfeld. The hidden costs of control. *American Economic Review*, 96:1611–1630, 2006.

E. Fehr and B. Rockenbach. Detrimental effects of sanctions on human altruism. *Nature*, 422:137–140, 2003.

U. Fischbacher. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178, 2007.

U. Fischbacher, E. Fehr, and S. Gächter. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13:1–25, 2002.

J.R.J. Fontaine, P. Luyten, P. de Boeck, J. Corveleyn, M. Fernandez, D. Herrera, A. Ittzes, and T. Comcsanyi. Untying the gordian knut of guilt and shame: the structure of guilt and shame reactions based on situation and person variation in Belgium, Hungary and Peru. *Journal of Cross-Cultural Psychology*, 37:273–93, 2006.

O. Frohlich and J. Oppenheimer. Some consequences of e-mail vs. face-to-face communication in experiment. *Journal of Economic Behavior & Organization*, 35:389–403, 1998.

U. Gneezy and A. Rustichini. A fine is a price. *Journal of Legal Studies*, 29: 1–17, 2000a.

U. Gneezy and A. Rustichini. Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115:791–810, 2000b.

B. Greiner. The online recruitment system orsee 2.0 - a guide for the organization of experiments in economics. *University of Cologne Working Papers*, 2004.

A. Hopfensitz and E. Reuben. The importance of emotions for the effectiveness of social punishment. Technical report, Tinbergen Institute, Discussion paper 05-075, 2007.

P.J. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In L.M. Le Cam and J. Neyman, editors, *Proceedings of the fifth Berkely symposium in mathematical statistics, and probability (Vol. 221-233)*. Prceoedings of the fifth Berkely symposium in mathematical statistics, and probability (Vol. 221-233), Berkeley and Los Angeles: University of California Press., 1967.

N.L. Kerr and C.M. Kaufman-Gilliland. Communication, commitment and coordination in social dilemmas. *Journal of Personality and Social Psychology*, 66:513–29, 1994.

C. Keser and F. van Winden. Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, 102:23–39, 2000.

T. Ketelaar and W.T. Au. The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: an affect-as-information interpretation of the role of emotion in social interaction. *Cognition & Emotion*, 17:429–53, 2003.

J.O. Ledyard. Public goods: a survey of experimental research. In J.H. Kagel and A.E. Roth, editors, *The handbook of experimental economics*, pages 111–94. Princeton University Press, New Jersey, 1995.

R. Lopez-Perez. Aversion to norm-breaking: a model. *Games and Economic Behavior*, forthcoming, 2008.

T. Miettinen. Promises and conventions. an approach to pre-play agreements. Discussion paper on strategic interaction, Max Planck Institute of Economics, 2006.

J.M. Orbell, A.J.C. van de Kragt, and R.M. Dawes. Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology*, 54: 811–9, 1988.

E. Ostrom. The value-added of laboratory experiments for the study of institutions and common-pool resources. *Journal of Economic Behavior & Organization*, 61:149–63, 2006.

M. Robinson and G. Clore. Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128:934–60, 2002.

D. Sally. Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992. *Rationality & Society*, 7:58–92, 1995.

R.H. Smith, J.M. Webster, W.G. Parrott, and H.L. Eyre. The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, 83:138–59, 2002.

J.P. Tangney, J. Stuewig, and D.J. Mashek. Moral emotions and moral behavior. *Annual Review of Psychology*, 58:345–72, 2007.

H. White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25, 1982.