Marco Faillo, Federico Fornasari, and Luigi Mittone

Tell Me How to Rule: Leadership, Delegation, and Voice in Cooperation

CEEL Working Paper 4-16

Cognitive and Experimental Economics Laboratory

Via Inama, 5 38100 Trento, Italy

http://www-ceel.economia.unitn.it
tel. +39.461.282313

Tell Me How to Rule: Leadership, Delegation, and Voice in Cooperation

Marco Faillo* Federico Fornasari† and Luigi Mittone‡

April 2016

Abstract

Following some recent studies, we experimentally test the effect of intra-group leadership in a public good experiment. Specifically, individuals taking part in our experiment are randomly assigned either the role of leader or the role of follower. Leaders take part in a public good game, aware of the fact that every decision they make directly affects their followers. In this sense, our experimental setting combines the dimension of leadership in cooperation with the one of delegated agents. In our experiment, we find that leadership produces two main effects: subjects contribute more, and tend to punish more frequently. In spite of the presence of higher contributions, we observe lower payoffs; these are caused by an aggressive behavior that push leaders to mane an undue use of punishment. Allowing one-sided communication between followers and leaders provide a different effect: communication reduces decision makers' aggressiveness, leading to lower contributions and punishment, but better results in terms of final payoffs. The same welfare can be reached when leadership is not implemented at all; this suggests that the presence of a dictatorial leader in public goods with punishment can be beneficial only when there is communication.

Keywords: Voluntary contribution experiment; Leadership; Punishment

JEL-classification: C72; C92; H41; O12

*DEM, University of Trento. E-mail: marco.faillo@unitn.it

†School of Social Sciences, University of Trento. E-mail: federico.fornasari@unitn.it

‡DEM, University of Trento. E-mail: luigi.mittone@unitn.it

1 Introduction

Cooperation in finitely repeated Public Good experiments has been widely tested during the past, producing results that, in general, have generated discrepancies from standard economic theories. According to these, subjects should never contribute to the Public Good, as this strategy does not represent an equilibrium. Experimental evidence (Fehr and Gächter, 2000; Chaudhuri, 2011) shows that the general tendency in Public Good games is for subjects to start contributing an amount within the 40 and 60% of their endowments and, as the game proceeds, to decrease contribution. One of the causes of such a decay is often attributed to the *end-game effect*: when approaching the end of a finite game, subjects tend to reason using backward induction and to free ride. Yet, an alternative theory is that subjects, during the game, go through a learning process and adapt their behavior, changing strategy (Isaac and Walker, 1988). such a decay in contribution is not observed when punishment is implemented (Fehr and Gächter, 2000).

More recently, studies involving public good experiments have driven the attention to leadership and its effects on cooperation tasks. Güth et al. (2007) proved that leadership affects intra-group cooperation. The authors implemented a leadership by example mechanism, where leaders are first movers and followers can imitate their choices: results show that the presence of a first mover-leader cause a substantial improve in cooperation.

A more recent study by Fleiß and Palan (2013) compares voluntary contributions and leader-allocated contributions in public goods. Results show that public-good games with allocators achieve higher levels of cooperation and contribution. Furthermore, the authors prove that the majority of the subjects taking part to their experiment is willing to delegate decisions to a leader in order to exploit benefits of cooperation. Such being the case, leaders could take advantage of their position to reap the benefits of cooperation, choosing higher contributions for their follower and deciding to free ride. In spite of this, the presence of a leader is usually beneficial for the economy of a group.

Hamman et al. (2011), similarly, explain how centralized decision-making is more efficient for the provision of a public-good. The authors find among subjects a general necessity for actions coordination while pursuing a common goal, especially in settings with large groups. Furthermore, the authors find that communication fosters coordination, by reducing the problem of free-riding.

General evidence supports the hypothesis that leadership improves cooperation in social dilemmas; nevertheless, we are not aware of the effects that leadership can have on cooperation game where leaders interact with each others. In our experimental study, leaders represent their own groups interacting into a public good game, but every choice they make directly falls back on their followers, that have no decisional power. This experimental setting differs from the ones used so far in this field, as we combine two different dimensions: leadership in coordination games, and delegated decision-making.

At the best of our knowledge, our experiment is the first aimed to investigate how delegated leaders, whose incentives are aligned to their followers', interact into a Public Good experiment. Most of the existing experiments are aimed to test how leadership affect intra-group cooperation, meaning that leaders are asked to interact only with their followers, leaving aside cooperation between leaders.

In particular, we focus on the implementation of leadership with dictating power, leaving aside the aspects of leadership by example (Güth et al., 2007; Levati et al., 2007). Our study is mainly aimed to understand whether leader-ship can help improve cooperation when individuals have to play a public good game, and they know that their actions directly affect others: in this sense, leaders are responsible for their own group and are asked to make delegated choices. As explained by Humphrey and Renner (2011), when delegated agents have to interact with other individuals, the sense of responsibility affects their choices, as they perceive their power to determine others' payoffs.

Similarly to what done in other experiments on cooperation in public goods experiments, we included the possibility for subjects to implement punishment (Fehr and Gächter, 2002; Egas and Riedl, 2005). Previous studies proved that, depending on the cost, subjects punish more or less frequently, and that, in general, punishment has a positive effect on cooperation: it is usually able to increase contributions and to reduce free riding.

Hamman et al. (2011) found that subjects taking part in a public good experiment show willingness to delegate their decisions only when communication is allowed. Xiao and Houser (2009) and Ellingsen and Johannesson (2008) provide evidence of the importance of one-sided communication in dictator games: when recipients have *voice*, this works as a psychological device that reduces dictators' aggressiveness. Additional contributions to this literature are provided by Capizzani et al. (2015) and Mittone and Musau (2016), who find similar results testing communication in social dilemmas. Following these findings, we decided to test the effect of communication in our experimental setting, by adding an additional treatment where one-sided communication is allowed at the end of each round.

Decisions made by subjects taking part to the treatments with simple leadership and leadership with communication were compared to the ones obtained in the public good game with punishment. The setting we implemented is similar to the one used by Fehr and Gächter (2000), that only differs in the number of rounds.

Our findings demonstrate that contribution to the public good is, on average, higher in the treatment with leadership. The same can be said about the use of punishment, which is implemented more frequently when leaders can decide and communication is not allowed. The higher frequency of punishment, though, jeopardizes the positive effects on efficiency derived from leadership, so that results in terms of final payoff are better in groups with no leaders (i.e. our baseline), and in groups with leadership and communication.

We can say that, in our experimental setting, leadership has, overall, a positive effect on contribution and cooperation. Nevertheless, communication is a needed in order to push leaders to make choices that are actually beneficial to their group. In addition to this, we find that implementing leadership with communication it is possible to obtain average final payoffs that do not significantly differ from the one obtained by subjects deciding individually.

2 Methodology

2.1 Task

The experiment is based on a repeated Public Good game with free contribution and punishment opportunities (Fehr and Gächter, 2000), and consists of three different treatments, applied in a between subjects fashion. In all the treatments participants are divided into groups of 4 members each, and every group plays a separate Public Good game over 20 rounds. Each round consists of two phases. During the first (contribution) phase, subjects receive an endowment E of 20 tokens, and have to choose how many of these they want to allocate to the Public Good. Every token invested in the Public Good is multiplied by 1.6 and then equally divided among the four members of the group. At the end of this first phase, subjects are informed about their payoffs. This is the sum of the tokens not allocated to the Public Good and a share equal to the 40% of the total sum invested by the group: Thus, subject i's payoff for the first phase is determined as follows:

$$\pi_{i,1} = (E - c_i) + \alpha * \sum_{i=1}^{4} c_i,$$

where c_i is subject i's contribution to the Public Good, and α is the Public Good multiplication factor equal to 0.4. Subjects' payoffs are all determined in the same way and every member of a group receives the same share of the Public Good.

After receiving information about their payoff, subjects enter the second phase of the round and receive details about other members' contributions. At this point, each group member can decide whether to reduce or leave unchanged the payoffs obtained by every other member during the first phase of the round. To do this, each subject can assign to his peers up to 10 points: every point that participants receive reduces by 10% the payoff accumulated in the first phase. Specifically, subjects can decide how many points they want to assign to each one of their group members. If subjects do not want to reduce others' payoff they have to assign 0 points. Else, they can decide by what percentage to reduce others' payoff and choose the corresponding number of points. To assign points is costly and the price varies depending on the number of points one wants to distribute. Within the experiment instructions, subjects are provided with a table that summarizes the cost of points.

Table 1: Points Cost

POINTS	0	1	2	3	4	5	6	7	8	9	10
COST	0	1	2	4	6	9	12	16	20	25	30

At the end of the second phase, subject i's payoff function is as follows:

$$\pi_{i,2} = \pi_{i,1} - \sum_{j \neq i} cost_{ij,2} - 0.10 * \pi_{i,1}(\sum_{j \neq i} p_{ji,2}),$$

where $\pi_{i,1}$ is subject i's payoff at the end of the first phase, $cost_{ij,2}$ is the cost for the points assigned by subject i and $p_{ji,2}$ are the points assigned to player i, both during the second phase.

By assigning 10 points to subjects, it is possible to reduce their first phase payoff by 100%. Following the experimental structure by Fehr and Gächter (2000), in order to prevent negative payoffs, even if subjects can receive more than 10 points, their payoff cannot be reduced by more than 100%. After having chosen how many points to assign, subjects receive feedback about this second phase and proceed to the following round. This procedure is repeated for 20 rounds, then subjects are informed about their final payment, that is equal to the sum of the payments obtained in the 20 rounds.

2.2 Treatments

The experiment consists of three different treatments: one serves as a baseline and is a replication of Fehr and Gächter (2000), while the other two include manipulations.

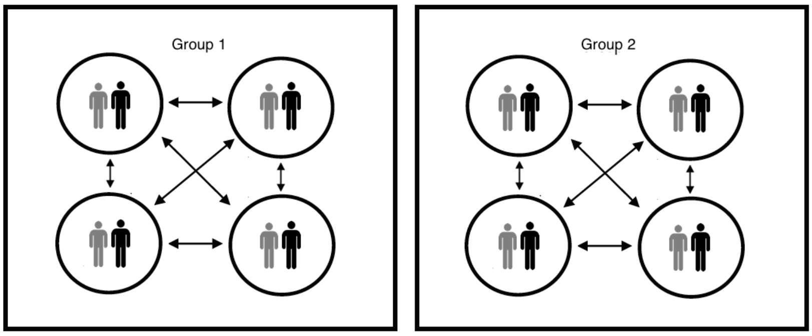
We use the *baseline* to gather data about individuals average contributions within the Public Good game described in the previous section. Average contributions are used to drive a comparison with the ones obtained in the two manipulation treatments.

The two manipulation treatments allow us to test two different factors: the first is how the awareness of being responsible for others affect individual decision making (treatment *couple*); and the second add to responsibility for others, also an higher social pressure as cheap-talk is allowed (treatment *chat*).

In the treatment *couple* subjects are matched in four couples; thus, every group is composed by eight members: four *drawn participants* and the four participants they are paired with (henceforth *followers*); the composition of couples and groups remain the same during the whole experiment.

Figure 1 represents a standard experimental session with 16 participants: these are divided into two groups of four couples, each consisting of a *leader* and a *follower*, represented by the black and the gray silhouette respectively. The arrows explain how *leaders* of couples belonging to the same group interact among each other, playing the Public Good game.

Figure 1: Group Structure - *Couple*



Each one of the 2 groups plays the 20 round Public Good game with punishment described in the previous section. Depending on the role participants are assigned to, they are asked to perform a different task: in the first phase of every round, *drawn participants* have to choose how much to contribute to the *common project*, and in the second phase they can decide how many punishment points to assign to other group members; *followers* can only observe what *drawn participants* do, but they are asked to express their hypothetical choices as if they were *drawn participants*. This helps to keep them busy and to keep private the identity of the *drawn participants* within the laboratory.

As *drawn participants* decide for themselves and the participants they are paired with, in the first phase their are given an endowment of 40 tokens (i.e. double with respect to the one provided in the *baseline*). Tokens invested in the *common project* are multiplied by 1.6 and equally divided among the 4 group members. Since in this treatment group members are, in fact, couples, participants receive half of the payoff addressed to their couple, which is entirely determined by the decisions made by the *drawn participants.* Once every *drawn participant* has chosen how much to contribute, the first phase ends. At the beginning of the second phase participants are informed about their group contribution and *drawn participants* can decide to assign some points to other group members. Similarly to what happens in the first phase, the costs of the points is doubled (Table 2). *Drawn participants* are entitled to decide how many points they want to assign, but the cost of this action is to be equally divided among the members of the couple.

Table 2: Points Cost - *Couple* Treatment

POINTS	0	1	2	3	4	5	6	7	8	9	10
COST	0	2	4	8	12	18	24	32	40	50	60

Note that if a couple receives points both members are affected and their payoffs decrease. After that all the *drawn participants* have decided how many points to assign, all the participants receive updated feedback about their payoff for the round and proceed. This procedure is repeated for 20 rounds, then subjects are informed about their final payment and the experiment is over.

The *chat* treatment is characterized by the same experimental setting utilized for the *couple* treatment, with the addition of one feature: followers are still asked to express their hypothetical decisions, but they can also send brief communications the *drawn participants* their are paired with. Communication of personal information, PC number, threats, promises of side payments and the use of offensive language were prohibited.

2.3 Participants and Procedures

The experiment was conducted in the Cognitive and Experimental Economics Laboratory (CEEL) of the University of Trento. The experiment was designed and administrated by using Borland Delphi.[1] Participants recruited were, on average, 23 years old second year students, 55% of them from the faculty of economics, and 45% are females. On average, participants had taken part to 5

[1]We express our sincere appreciation for Marco Tecilla's support.

experiments before. The total number of subjects is 360, and they were divided across treatments as shown in Table 3.

Table 3: Number of Participants

Treatment	Sessions (Num.)	Participants (Num.)	Observations (Num.)
baseline	4	72	18
couple	10	160	20
chat	8	128	16
Total	22	360	54

Each subject received a 3.00 euros show-up fee, plus a sum that varied depending on their performance in the experiment; this was, on average, equal to 8 euros. Upon their arrival, subjects were randomly assigned to a computer and received instructions for the experiment, which also contain few comprehension questions; subjects had 7 minutes to read the instructions and try to answer. Then, instructions were read aloud and the correct answer were provided by one of the experimenters, who also answered any possible question.

2.4 Behavioral Predictions

With regard to the *c*ouple and the *c*hat treatment, Charness and Jackson (2009) report experimental evidence of the effects that being responsible for someone's payoff has on subjects. Specifically, the authors tested dictators' leadership in a coordination game, finding that the majority of subjects showed a greater risk aversion when responsible for others. This study have extended the *responsibility-alleviation effect* (Charness, 2000) to the dimension of coordination games, showing how the presence of a unique decision maker in a two-persons group can work as an instrument to increase coordination and welfare.

Additional evidence about leadership and cooperation in public good games can be found is some recent experiments (among the others Fleiß and Palan, 2013; Bolle and Vogel, 2011). These, although have diverse experimental settings and manipulations, are characterized by a common finding: leadership improves cooperation.

With respect to the aspects described, and accordingly to our premises, we formulate the following behavioral predictions:

Hypothesis 1 *Contributions:*
In all the three treatments participants will contribute to the Public Good.

Hypothesis 2 *Social Welfare:*
When representing a couple, subjects will invest more in the Public Good.

Hypothesis 3 *Punishment:*
Subjects will punish free-riders and low contributors; responsibility for others increases the use of punishment.

Hypothesis 4 *Communication:*
In the treatment chat followers will communicate with leaders affecting their decisions.

3 Results

This section includes an analysis divided into three main focus areas: contributions, payoff and punishment. Note that, with regard to the treatments *couple* and *chat*, we only use decisions made by individuals assigned the role of *drawn participants*; in fact, as explained in the Methodology section, *followers* were asked to express their decisions with the only aim of keeping them busy during the experiment and preserving roles anonymity. In addition to this, we include a section specifically addressed to the analysis of communication in treatment *chat*.

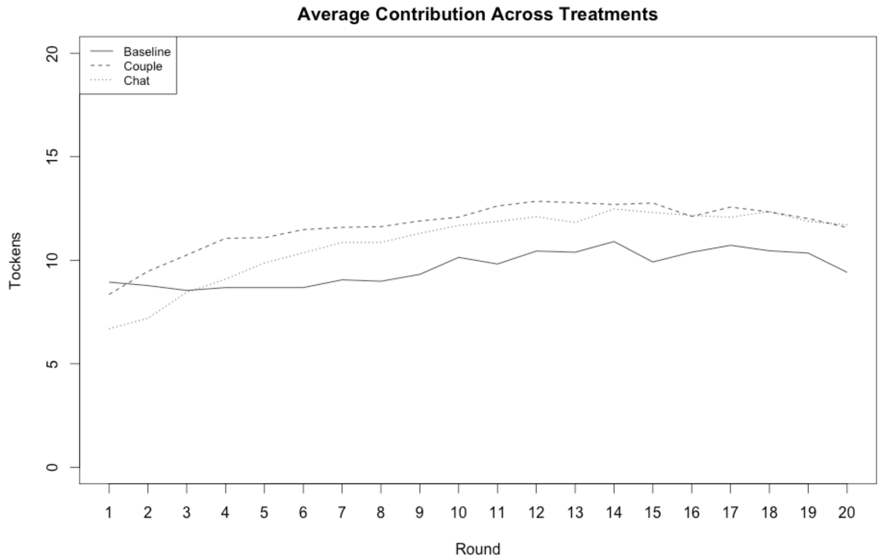
3.1 Contributions

Figure 2 shows the average percentage contribution to the Public Good per round across the three treatments.

We observe both *couple* and *chat* treatments are characterized by higher average contributions with respect to the *baseline*.

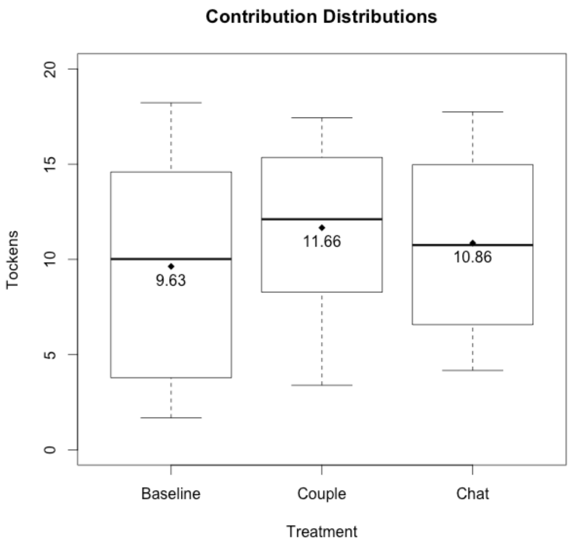
Result 1 *In all the three treatments subjects contribute to the Public Good investing positive sums.*

Figure 2: Average Percentage Contribution per Round



The boxplot in Figure 3 provides more precise information about the difference in the average percentage contributions across treatments.

Figure 3: Average Percentage Contribution per Treatment



The three distributions seem to confirm the presence of a difference in terms of contributions between the *baseline* and the other treatments. In the *baseline* subjects have contributed on average with 48% of their endowment, while in the *couple* and *chat* treatments average individual contributions are, respectively, equal to 58% and 54% (black dots and numbers). Non-parametric tests show a significant difference in contributions between the *baseline* and *couple* treatments.[2]

Result 2 *In all the three treatments contributions to the Public Good will be, on average, positive.*

Result 3 *When deciding for others (and no communication is allowed), subjects tend to invest more in the Public Good trying to increase their group welfare.*

The difference between *baseline* and *chat* is not significant according to non-parametric tests.[3]

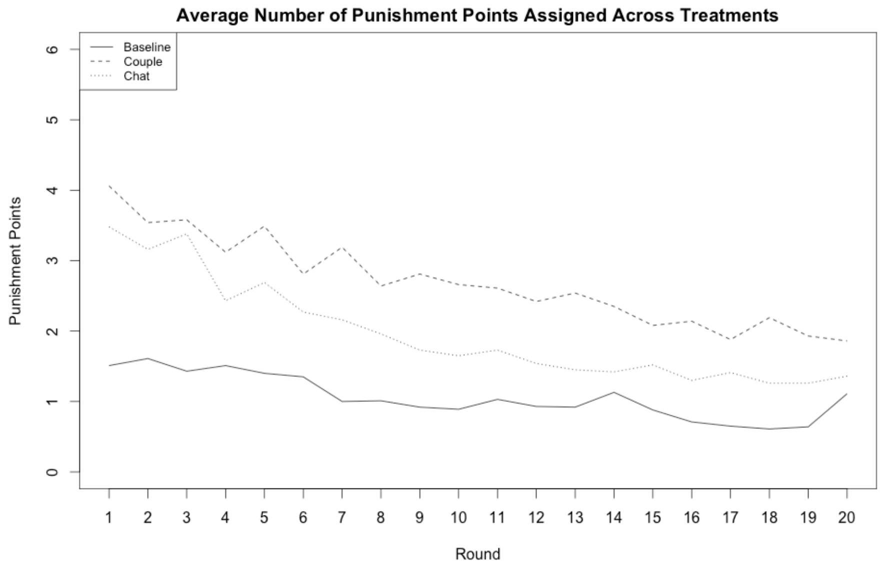
[2]One-tailed tests on groups' average values across rounds: Wilcoxon-Mann-Whitney, p-value = 0.0834; Fligner-Policello, p-value = 0.0778.

[3]Nevertheless, figure 2 suggests that the first five rounds could be considered a transition phase: *couple* contributions start around the 42%, while *chat* contributions start at 33% and both quickly grow by 10 and 15 percentage points. Thus, we tried to repeat our tests excluding data from these rounds. Results confirm our hypothesis; in fact, we find stronger evidence

3.2 Punishment

As for the level of punishment (figure 4), we find strictly positive values in every treatment.

Figure 4: Average Points of Punishment per Round



In general, punishment seems to be implemented more frequently in treatments with delegation, that do no significantly differ from each other[4]; this suggests that subjects punish more when responsible for others.

As already pointed out in previous studies (Herrmann et al., 2008), it is possible to make a distinction between two different types of punishment: the *altruistic punishment* that aims to punish free-riders and low cooperators, and the so called *antisocial punishment* directed to cooperators.[5] We found evidence of the implementation of both these forms of punishment.

of the difference between *baseline* and *couple* (One-tailed tests: Wilcoxon-Mann-Whitney, p-value = 0.0739; Fligner-Policello, p-value = 0.0697), but results are still negative for as concerns the *chat* treatment.

[4]Two-sidedWilcoxon-Mann-Whitney on group average across rounds: *couple* vs *baseline* p-value = 0.0006; *chat* vs *baseline* p-value = 0.007; *couple* vs *chat* p-value = 0.13

[5]Our experimental design is not aimed to investigate and provide an explanation to this particular phenomenon; thus, we leave a deeper analysis to further studies.

Table 4: Average Points of Punishment Assigned

Treatment	**Points Assigned**	**Antisocial (%)**
baseline	1,06	21,2%
couple	2,69	18,2%
chat	1,76	14,5%

Points Assigned: average number of punishment points assigned; *Antisocial (%)*: percentage of antisocial punishment over the points assigned.

Values in table 4 represent the average number of points of punishment assigned across treatments. In general, it is possible to observe how treatments *couple* and *chat* are characterized by a greater use of punishment. As it was already pointed out by the results of the non-parametric test, leaders punished more than subjects in the *baseline*.

Result 4 *Subjects punish free-riders and low contributors. Leadership increases the use of punishment.*

For as concerns antisocial punishment, our results are in line with previous studies, but we observe a difference across the three treatments. Running a non-parametric test we find that antisocial punishment in the treatment chat is lower than in the *baseline*; furthermore, we observe that in the treatment *couple* antisocial punishment is higher than in the *chat* one.[6]

3.3 Payoff

Our results suggest, so far, that decision makers in charge as leaders tend to contribute more to the common project and also to punish more. Nevertheless, we also find that these effects are weakened in the *chat* treatment. Thus, in order to have a more clear understanding of subjects' welfare, it is more appropriate to focus on payoffs.

[6]Two-sided Wilcoxon-Mann-Whitney on groups' average: *chat* vs *baseline* p-value = 0.09; *chat* vs *couple* p-value = 0.05.

Figure 5: Average Payoff per Round

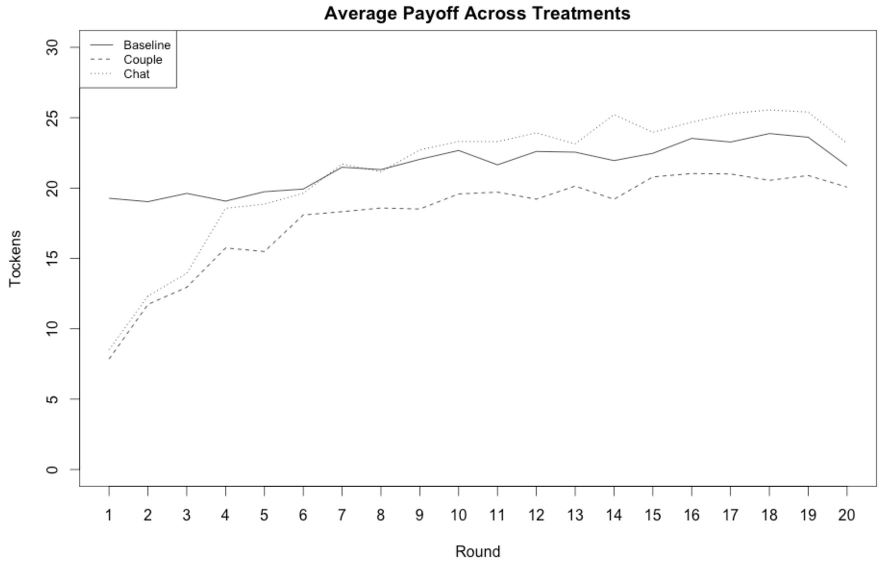
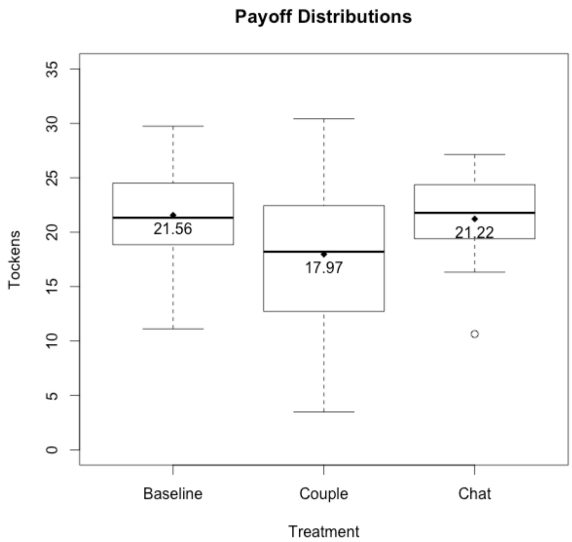


Figure 5 reports the average payoff at the end of every round; this means that payoffs are taken after that punishment has occurred. Contrarily to what found concerning contributions, here we observe that payoffs in the treatment *chat* are higher than in the treatment *couple*. This suggests that, in spite of the lower level of contributions, subjects decided to punish less in order to preserve general welfare. Furthermore, payoffs in the *baseline* follow a flatter pattern, but their level does not seem to differ from the other treatments. More details are provided by the following box plot (figure 6).

Figure 6: Average Payoff per Treatment



The mean in the treatment *chat* is higher than in the treatment *couple*, but what is more interesting is that the highest mean value is associated to the *baseline*. As already explained, responsibility for others has the effect of increasing efficiency, thus we can use a one-tailed Wilcoxon-Mann-Whitney test on groups' average values across rounds, which confirms that the individual average payoff is, indeed, higher in treatments *baseline* (p-value = 0.0559) and *chat* (p-value = 0.0580) rather than in *couple*.

This finding is particularly relevant as it suggests that, in our experimental setting, leadership can be used to enhance social welfare only in presence of communication; in fact, in spite of having contributed more on average, subjects in the *couple* treatment had up a significant part of their group's welfare in order to actively sustain cooperation.

3.4 Chat

In the treatment *chat*, *followers* were allowed to send brief messages to the *drawn participant* they were paired to. 67.19% of the 64 *followers* has used the chat at least once during the 20 rounds, for a total number of messages equal to 554. We have categorized these into 5 types of communication: suggestions to increase or decrease the contribution, suggestions to increase or decrease the punishment, and other messages. In addition to this, we checked whether the message was effective by monitoring *drawn participants*' behavior in the following round; we decided to leave aside messages of the category *other*, as their efficacy cannot be categorized and their content is not aimed to influence leaders' decisions. Table 5 provides data of the analysis.

Table 5: Chat

	Effective		
Message	**No**	**Yes**	**Total**
increase contribution	28	97	125
decrease contribution	21	63	84
increase punishment	7	12	19
decrease punishment	8	31	39
Total	64	203	267

In details, 36% of the total number of messages analyzed in table 5 suggested, successfully, to increase the contribution (against 10% suggesting a decrease). The number of messages related to punishment is, in general, lower, but it is still possible to observe how the 24% of those led to a decrease in punishment.[7]

The overall effectiveness rate is equal to 76%, which suggests that *drawn participants'* actions were partly driven by their peers' suggestions. This is particularly relevant as it seems, looking at our results, that the use of chat helped manage resources more carefully; in fact, in the treatment *couple* we do observe higher contribution, but the use of punishment is so high that it jeopardizes any benefit.

Result 5 Followers *communicate using the chat and this affects the decision maker.*

3.5 Regressions

Table 6 presents three models the determinants of the dependent variable *individual contribution* of two models obtained from a Arellano Bond regression: the use of this model is required by the dynamic nature of our experimental data.

Model 1 contains two dummy variables, *treatment couple* and *treatment chat* (respectively referred to the treatments we applied), a time variable, *round*, and five lagged variables: *contribution t-1* is the individual contribution, *group average contrib t-1* is the group average contribution, *given punishment t-1* and

[7]We ran Z-tests on proportions to compare, respectively, messages suggesting to increase or decrease contribution, and messages suggesting to increase or decrease punishment: both tests rejected the null hypothesis.

received punishment t-1 are the punishments given and received, and *group antisocial t-1* is the group average antisocial punishment.

Model 2 has the same body, but we added four control variables: *gender* and *age* are demographic information, *major* is a dummy variable equal to 1 if the subject is a student of the faculty of Economics,and *experience* is the number of experiments at which the subjects had taken part.

Treatment Chat only refers to leaders' contribution in the treatment with communication, and includes two dummy variables: *chat increase contrib t-1* indicates that the leader, at the end of the previous round, has received a message by her follower, suggesting to increase their couple's contribution; *chat decrease contrib t-1* indicates that the leader, at the end of the previous round, has received a message by her follower, suggesting to decrease their couple's contribution. These variables are equal to one only when messages were successful, i.e. when leaders adapted their contributions according to the messages received.[8]

[8] See section *Chat* for more details.

Table 6: Leader's Contribution Determinants - Arellano Bond Model

	Model 1	Model 2	Treatment Chat
Individual Contribution			
treatment couple	0.277 (0.11)*	0.220 (0.12)	
treatment chat	0.262 (0.11)*	0.491 (0.12)*	
round	−0.026 (0.01)**	−0.023 (0.01)*	−0.036 (0.02)*
contribution t-1	0.194 (0.02)***	0.188 (0.02)***	0.405 (0.03)*
group average contrib t-1	0.752 (0.02)***	0.755 (0.02)***	0.546 (0.03)***
given punishment t-1	−0.009 (0.02)	0.000 (0.02)	−0.038 (0.03)
received punishment t-1	0.060 (0.03)*	0.064 (0.03)*	−0.028 (0.05)
group antisocial t-1	−0.006 (0.00)	−0.006 (0.00)	0.011 (0.01)
gender		0.100 (0.09)	−0.184 (0.15)
age		−0.028 (0.02)	−0.060 (0.04)
major		−0.332 (0.09)***	−0.243 (0.17)
experience		0.064 (0.01)***	0.081 (0.03)
chat increase contrib t-1			3.180 (0.30)***
chat decrease contrib t-1			−2.969 (0.36)***
cons	0.820 (0.15)***	1.225 (0.48)*	2.455 (0.98)*
Number of obs.	4104	4104	1216
Wald Chi-sq	15146.18	15324.55	6145.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

treatment couple and *treatment chat* are dummies indicating the treatment applied, *contribution t-1* is the individual contribution in the previous round, *group average contrib t-1* is average contribution by group in the previous round, *given punishment t-1* is the punishment subject used in the previous round, *received punishment t-1* is the punishment subject received in the previous round, *group antisocial t-1* is the average antisocial punishment used in the previous round by the group, *experience* is the number of experiments the subjects took part to, *chat increase contrib t-1* and *chat decrease contrib t-1* indicates whether leaders received a successful message suggesting, respectively, to increase or decrease their contribution.

In general, our models seem not to differ widely both in terms of parameters and their significance. The treatments dummy variables are both significant in Model 1, supporting the idea that leadership has the effect to increase contribution in our experimental setting. In Model 2, the dummy related to the treatment *couple* loses significance: this may be caused by the larger number of independent variables included.

The variable *round* has a negative coefficient in all the models, and this seems to conflict with the trend observed in figure 2; nevertheless, as the lagged individual contribution variable is always positive, the presence of a negative parameter related to the variable *round* suggests that there is a flattening in contributions as the game proceeds.

The average group contribution in the previous round positively affect individual contributions, so that, to some extent, we could say that cooperation sustains itself. Similarly, in models 1 and 2 punishment received in the previous round has a positive effect on individual contribution. This effect is not observed in the treatment *Chat*: an explanation can be found in the high significance of the coefficient related to the variable *chat decrease contrib t-1*. This suggests that followers' messages were taken into account by leaders, even when causing the couple to incur punishment.

In addition to this, we find that also the coefficient of the other dummy variable related to chat, *chat increase contrib t-1*, is highly significant, supporting our result that points out the efficacy of the use of communication between followers and leaders.

For as regards our control variables, we only observe significant effects in Model 2: Economics students tend to contribute less than others, while people with more experience in laboratory experiments contribute more to the Public Good.

4 Conclusion

Following some recent work aimed to study leadership and cooperation (Hamman et al., 2011; Bolle and Vogel, 2011; Güth et al., 2007), we test in a laboratory experiment the effects of leadership on intra-group cooperation, com-bining this condition with delegated choices. Specifically, we investigate whether leadership is beneficial when, in addition of being responsible for others, leaders are also asked to cooperate among themselves. This aspect is particularly inter-esting and, at the best of our knowledge, have not been tested experimentally yet. Furthermore, we compare a situation where leaders decide independently to another where their followers are allowed to send short messages.

Our main finding is that, when there is no communication, leaders tend to contribute more to the public good, but also to make an undue use of punishment; the latter aspect is true when referring to both *altruistic punishment* and *antisocial punishment*. When *followers* are allowed to communicate with their leaders, providing their opinion about the choices made in the round just concluded, it is possible to observe a decrease in contributions and punishment, such that they do not significantly differ from the average ones observed in our baseline.

In terms of final payoffs, although *leaders* contribute and cooperate more in the treatment *couple*, we observe, on average, higher earnings in the baseline and in the treatment *chat*. This result suggests that the responsibility for others that *leaders* perceive when contributing to determine others' payoffs incentivizes them to cooperate more, but also to punish more, with the aim of preventing other leaders from compromising their effort. Nevertheless, punishment is used without considering the consequences, so that the effects of higher contributions are jeopardized.

From this perspective, it is possible to reduce punishment[9] and obtain higher final payoffs providing followers with the possibility of communicating with leaders they are matched with. We find that results from the treatment with leadership and communication produces very similar results to the public good game with punishment we use as baseline.

This evidence suggests that leadership itself may be, in some cases, more harmful rather than beneficial; in fact, decision makers, moved by their sense of responsibility for others', seem not to be always able to correctly decide. When followers have *voice*, leaders tend to be less aggressive because communication works as a trigger that relieves the psychological pressure of responsibility (Xiao and Houser, 2009; Ellingsen and Johannesson, 2008): this, in our case, reduces the use of punishment and produces an increase in efficiency.

To summarize, we can conclude that, in our experimental setting, a normative leadership where leaders have dictatorial power increases, on average, contributions to the public good, but, because of leaders' aggressiveness in the use of punishment, does not produce beneficial effects on final payoffs. On the other hand, implementing a "weakened dictatorship" where communication is allowed, it is possible to obtain meaningful results in terms of both cooperation and efficiency, preventing waste of resources deriving from individual aggressiveness.

Our experimental results are interesting as they contribute to an unexplored dimension, shedding light on leaders' interaction and responsibility for others in a cooperation game. Nevertheless, it is important to keep in mind that we only provide a preliminary analysis. Further experiments should aim to investigate the interaction between leaders in order to provide useful insights and find how a greater cooperation could be achieved: this aspect could be relevant also for purposes regarding external validity.

[9] As explained in the section *Punishment*, the treatment *chat* does not significantly differ from the treatment *couple*; yet, as observed in figure 4, the number of points assigned seems to be smaller.

References

Bolle, F. and Vogel, C. (2011). Power comes with responsibility—or does it? *Public Choice*, 148(3-4):459–470.

Capizzani, M., Mittone, L., Musau, A., and Vaccaro, A. (2015). Anticipated communication in the ultimatum game. *working paper*.

Charness, G. (2000). Responsibility and effort in an experimental labor market. *Journal of Economic Behavior & Organization*, 42(3):375–384.

Charness, G. and Jackson, M. O. (2009). The role of responsibility in strategic risk-taking. *Journal of Economic Behavior & Organization*, 69(3):241–247.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83.

Egas, M. and Riedl, A. (2005). The economics of altruistic punishment and the demise of cooperation.

Ellingsen, T. and Johannesson, M. (2008). Anticipated verbal feedback induces altruistic behavior. *Evolution and Human Behavior*, 29(2):100–105.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, pages 980–994.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Fleiß, J. and Palan, S. (2013). Of coordinators and dictators: A public goods experiment. *Games*, 4(4):584–607.

Güth, W., Levati, M. V., Sutter, M., and Van Der Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, 91(5):1023–1042.

Hamman, J. R., Weber, R. A., and Woon, J. (2011). An experimental investigation of electoral delegation and the provision of public goods. *American Journal of Political Science*, 55(4):738–752.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Humphrey, S. J. and Renner, E. (2011). The social costs of responsibility. *CeDEx discussion paper series*.

Isaac, R. M. and Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, pages 179–199.

Levati, M. V., Sutter, M., and Van der Heijden, E. (2007). Leading by example in a public goods experiment with heterogeneity and incomplete information. *Journal of Conflict Resolution*, 51(5):793–818.

Mittone, L. and Musau, A. (2016). Communication, sequentiality and strategic power. a prisoners? dilemma experiment. Technical report, Cognitive and Experimental Economics Laboratory, Department of Economics, University of Trento, Italia.

Xiao, E. and Houser, D. (2009). Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange. *Journal of Economic Psychology*, 30(3):393–404.

A Experiment Instructions

Following we include an English translation of the experiment instructions. Our experimental design required us to produce three different version of the instructions (i.e. one for each treatment). General instructions were common, while instructions for the remaining part of the experiment were edited to match the structure of our treatments.

As explained in the section on the experimental task, the main differences in our experiment occur between the *baseline* and the treatments *couple* and *chat*. These latter, in fact, do not differ much from each other.

What follows is a full version of the instructions we used for the experimental sessions. Any time there will be an edited part, it will be noticed, specifying to which treatment we are referring (*baseline*, *couple*, or *chat*). May the instructions be common to all the treatments, this will indicated as well (*common*).

General Instructions

[*common*] Good morning, thanks for having accepted to participate to this experiment. You are taking part to a study on decisions in the economic environment. During the experiment you will have the opportunity to gain money. At this sum we will add 3 euros for your participation. Your payment will depends on your decisions and on other participants' decisions. The answers you give and the choices you make will be absolutely anonymous. The experimenters will not be able to associate your choices and your answers to your name. During the whole experiment we kindly ask you not to communicate with the other participants (otherwise, you will be excluded from the experiment) and to pay attention to the instructions that will be shown on your monitor and will be read aloud by one of the experimenter. May you have any question, ask the experimenters.

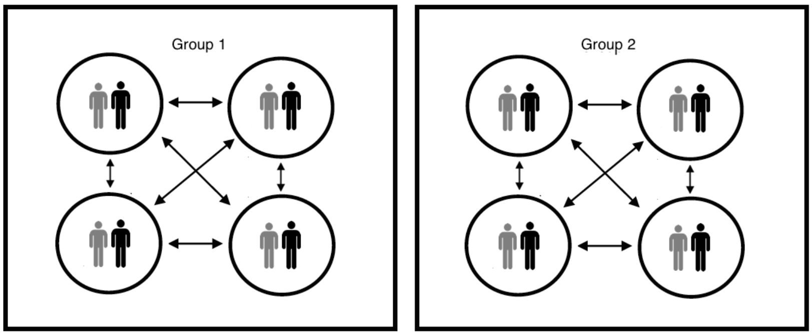
[*common*] Your payoff will be calculated in tokens: each token corresponds to 2 euro cents. At the end of the experiment we will ask you to fill a short questionnaire and we will proceed to the payment, that will be made in cash.

THE PARTICIPANTS

[*baseline*] Participants to the experiment are divided into 4 groups. The composition of groups is the same during the whole experiment. Thus, your group will be composed, in addition to you, by other three persons, whose identity you will not know.

[*couple*, *chat*] This experiment has 16 participants, divided into 8 couples. The composition of couples remains the same during the whole experiment. Thus, your couple will be made by you and another person, whose identity will remain unknown to you. In every couple, one of the participants will be randomly assigned the role of *drawn participant*, and the other to the role of *follower*. The *drawn participant* in your couple will interact with the *drawn participants* of other three couples according to the procedure described below. Your couple belongs to a group consisting of four couples. These groups are two, and their composition remains unchanged during the whole experiment (Figure 7).

Figure 7: Group Structure



[*common*] The experiment is divided into 20 rounds. In each round the other member of the group will not be identified by any name, so that their choices cannot be identified either.

THE PHASES

Each round is composed of **2 phases**.

Phase 1

[*baseline*] At the beginning of each round every participants receives 20 tokens. We will call this sum "endowment". You will have to decide how to use your endowment. In particular you will have to decide ow many of the 20 tokens to utilize to contribute to a project and how many you of them you want to keep for yourself. The other members of the group will have to do the same.

The tokens invested in the project will be multiplied by a yield of 1.6 and equally divided between the 4 members of the group.

At the end of phase 1, you will be given information about your earning, that consists of two elements:

A The part of the initial 20 tokens that you decided to keep for yourself (meaning 20 tokens minus the contribution to the project);

B Your payment deriving from the project, which is equal to 40% of the sum of the four members' contributions.

Then, your earning at the end of phase 1 is computed by the computer as follows:

Your earnings at the end of phase 1 =
(20 tokens - contribution to the project) +
40% * (members' total contribution to the project)

[*couple*, *chat*] At the beginning of each round every participants receives 40 tokens. We will call this sum "endowment". The drawn participant of each couple has to decide how to use your endowment. In particular she will have to decide ow many of the 40 tokens to utilize to contribute to a project and how many you of them you want to keep for yourself. The other members of the group will have to do the same.

The tokens invested in the project will be multiplied by a yield of 1.6 and equally divided between the 4 couples of the group.

At the end of phase 1, you will be given information about your couple's earning, that consists of two elements:

A The part of the initial 40 tokens that the drawn participant decided to keep for your couple (meaning 40 tokens minus the contribution to the project);

B Your payment deriving from the project, which is equal to 40% of the sum of the four couples' contributions in your group.

Then, your earning at the end of phase 1 is computed by the computer as follows:

Your couple earnings at the end of phase 1 =
(40 tokens - contribution to the project) +
40% * (couples' total contribution to the project)

[*baseline*] Each group member's earnings are computed in the same way; furthermore, each couple receives the same payment from the project.

[*couple*, *chat*] Each couple's earnings are computed in the same way; furthermore, each participant receives the same payment from the project.

[*baseline*] Imagine, for instance, that in your group one member contributes with 10 tokens, another contributes with 8 tokens, a third member contributes with 12 tokens and you decide to contribute with 10 tokens. The total group contribution is then 40 tokens. So, each member of the group receives from the project a sum equal to 40% of 40 tokens = 16 tokens. The earnings for the 4 members will be:

- first participant: $20 - 10 + 16 = 26$
- second participant: $20 - 8 + 16 = 28$
- third participant: $20 - 12 + 16 = 24$
- fourth participants: $20 - 10 + 16 = 26$

[*couple*, *chat*] Imagine that, for instance, in your group, the *drawn participant* of the first couple contributes with 10 tokens, the second couple's one contributes with 8 tokens, the third couple's one contributes with 12 tokens and your couple's *drawn participant* decides to contribute with 10 tokens. The total group contribution is then 40 tokens. So, each couple of the group receives from the project a sum equal to 40% of 40 tokens = 16 tokens. The earnings for the 4 couples will be:

- first couple: $40 - 10 + 16 = 46$
- second couple: $40 - 8 + 16 = 48$
- third couple: $40 - 12 + 16 = 44$
- fourth couple: $40 - 10 + 16 = 46$

[*common*] The software will always display the number of the current round and the earnings accumulated until that moment.

Phase 2

[*baseline*] At the beginning of Phase 2 you can observe how much the other group members have contributed to the project. At this point each member of the group can decide to reduce or leave unvaried other members' phase 1 earnings, by assigning points, up to a maximum of 10 points. Each of the point assigned reduces by 10% the phase 1 earnings of the participant who receives it. Thus, if you decide to assign 0 points to another group member, you will not modify that participant's earnings. If you assign 1 point you will reduce that participant's earnings by 10%. If a person receives in total 4 points, her phase 1 earnings will be reduced by 40%. If the person receives 10 or more points her earnings will be reduced by 100%. Assigning points has a cost, which depends on the number of points you decide to assign.

The table shows the correspondence between the number of points assigned and the cost to pay.

POINTS	0	1	2	3	4	5	6	7	8	9	10
COST	0	1	2	4	6	9	12	16	20	25	30

For instance, if you decide to assign 1 point to a member of your group and 3 points to a second member, the earning of the first person will be reduced by 10% while the earning of the second will be reduced by 30%. The cost you have to pay in total is equal to 1 token (for 1 point) + 4 tokens (for 3 points).

Your earning at the end of Phase 2 is computed as follows:

Your earnings at the end of Phase 2 =
earning at the end of Phase 1 -
cost of points assigned in Phase 2 -
0.10 * (points received in Phase 2) *
(earning at the end of Phase 1)

Note that if you receive 10 points, your earning in Phase 1 will be reduced by 100%. The maximum reduction is, anyways, equal to 100%, also in the case you received more than 10 points. Note also that at the end of Phase 2 your earning can be negative. This happens when the cost of the points you have decided to assign is higher than your earning. If you pay attention it will not be hard to avoid this.

[*couple*, *chat*] At the beginning of Phase 2 you can observe how much the other couples have contributed to the project. At this point each *drawn participant* can decide to reduce or leave unvaried other couples' phase 1 earnings, by assigning points, up to a maximum of 10 points. Each of the point assigned reduces by 10% the phase 1 earnings of the couple that receives it. Thus, if the *drawn participant* decides to assign 0 points to another group member, she will not modify that participant's earnings. If she assigns 1 point she will reduce that participant's earnings by 10%. If a couple receives in total 4 points, its phase 1 earnings will be reduced by 40%. If the couple receives 10 or more points its earnings will be reduced by 100%. Assigning points has a cost, which depends on the number of points the *drawn participant* decides to assign.

The table shows the correspondence between the number of points assigned and the cost to pay.

POINTS	0	1	2	3	4	5	6	7	8	9	10
COST	0	2	4	8	12	18	24	32	40	50	60

For instance, if the *drawn participant* decides to assign 1 point to a couple of your group and 3 points to a second couple, the earnings of the first couple will be reduced by 10% while the earnings of the second will be reduced by 30%. The cost your couple has to pay in total is equal to 2 token (for 1 point) + 8 tokens (for 3 points).

Your couple's earnings at the end of Phase 2 are computed as follows:

Your couple's earnings at the end of Phase 2 =
earning at the end of Phase 1 -
cost of points assigned in Phase 2 -
0.10 * (points received in Phase 2) *
(earning at the end of Phase 1)

Note that if your couple receives 10 points, Phase 1 earnings will be reduced by 100%. The maximum reduction is, anyways, equal to 100%, also in the case your couple received more than 10 points. Note also that at the end of Phase 2 your couple's earnings can be negative. This happens when the cost of the points the *drawn participant* have decided to assign is higher than your couple's earnings. If you pay attention it will not be hard to avoid this.

[*common*]Any negative payoff at the end of the experiment will be balanced by using your show-up fee, which will then be reduced by an amount equal to the suffered loss. May the loss be higher than the show-up fee your payment at the end of the experiment will be equal to zero.

[*couple*, *chat*] THE NOT-DRAWN PARTICIPANTS [*couple*, *chat*] During the experiment the members of the couples who have not been drawn as *drawn participants* will observe the choices of their *drawn participant* and will obtain information about the other couples' contribution level and about assigned and received points. In addition to this, during every round they will be asked to express their choices as if they where *drawn participants*, both in terms of contribution and points giving.

[*chat*] Furthermore, at the end of every round they will have the opportunity to send a message to the *drawn participants*. The message will have to be inherent to the experiment activity, and shall not contain information that may reveal the sender's identity, nor offensive and rude statements.

FINAL PAYMENT

[*baseline*] At the end of the experiment you will be informed about your total payment, to which it is added the 3 euros participation fee.

[*couple*, *chat*] At the end of the experiment you will be informed about your couple's total payment. Your personal payment will correspond to half of that sum, to which it is added the 3 euros participation fee.

CONTROL QUESTIONS

[*baseline*]

1. Participants are divided into groups of .... people each.
2. The composition of groups remains the same during the whole experiment, so you will always interact with the same people. [ ] True [ ] False
3. Phase 1: your contribution is equal to 10 tokens and other group members' contributions are: 10, 5, 0. Your earning will be: ....
4. If your earning in Phase 1 is equal to 20 tokens and you receive 5 points, your earning will be then equal to: ....
5. If you assign to the other three members of your group the following points: 2, 3, 5, the total cost of the points will be equal to: ....

[*couple*, *chat*]

1. Participants are divided into .... couples and .... groups.
2. Your couple interacts with other .... couples.

3. The composition of groups remains the same during the whole experiment, so your couple will always interact with the same couples. [ ] True [ ] False

4. Phase 1: your couple *drawn participant*'s contribution is equal to 10 tokens and other couple *drawn participants*' contributions are: 20, 10, 0. Your couple's earning will be: ....

5. If your couple's earning in Phase 1 is equal to 30 tokens and you receive 5 points, your couple's earning will be then equal to: ....

6. If the *drawn participant* assigns to the other three couples of your group the following points: 2, 3, 5, the total cost of the points will be equal to: ....

B Data

Following, we provide mean values of the most relevant variables used for our analysis (Table 7). Data are organized by treatment and by group.

Table 7: Individual Average Values per Group

Treatment	Group	Contribution	contri_sd	Punish	punish_sd	Payoff	pay_sd
baseline	1	16.688	3.922	.500	.827	27.740	3.535
baseline	2	18.225	3.114	.275	.616	29.741	2.594
baseline	3	3.775	3.822	.500	.914	20.544	3.135
baseline	4	3.375	3.491	.163	.462	21.454	3.272
baseline	5	5.975	2.199	.313	.587	22.488	2.375
baseline	6	1.675	3.133	.663	1.330	18.852	3.122
baseline	7	12.113	2.506	.475	1.055	25.798	2.521
baseline	8	15.275	4.503	2.800	2.291	17.570	3.634
baseline	9	14.013	5.290	2.138	2.448	19.1753	9.737
baseline	10	14.588	4.995	2.975	2.873	16.656	8.186
baseline	11	2.213	3.883	.913	1.995	18.028	3.949
baseline	12	3.788	3.828	.288	.7826	21.202	3.399
baseline	13	15.988	4.295	.513	1.043	27.179	4.973
baseline	14	7.100	3.407	3.463	3.987	11.119	5.362
baseline	15	9.913	5.171	.363	.917	24.511	4.578
baseline	16	8.375	6.097	.600	1.308	22.519	4.239
baseline	17	10.150	2.141	.438	.869	24.419	2.936
baseline	18	10.125	5.931	1.738	2.249	19.097	5.325
couple	19	3.263	3.265	.375	.848	20.666	2.391
couple	20	15.688	3.360	.238	.601	28.421	2.963
couple	21	17.625	3.895	.038	.191	30.413	3.213
couple	22	10.750	8.511	4.225	4.466	9.577	15.430
couple	23	7.706	1.950	1.475	1.534	19.526	3.542
couple	24	15.588	5.539	1.650	1.700	22.678	5.779
couple	25	18.463	4.612	.813	2.007	27.266	7.559
couple	26	18.875	3.643	.400	1.481	29.584	4.877
couple	27	10.994	3.820	2.913	3.273	15.215	7.223
couple	28	8.856	5.423	3.088	3.562	13.183	8.459
couple	29	7.250	5.426	3.200	3.395	12.089	9.053
couple	30	9.988	1.859	1.475	2.289	20.524	5.090
couple	31	18.888	3.163	2.038	2.957	22.200	12.079
couple	32	9.725	7.382	2.100	3.197	17.541	10.593
couple	33	5.575	1.553	1.275	1.902	18.874	4.234

Treatment	Group	Contribution	contri_sd	Punish	punish_sd	Payoff	pay_sd
couple	34	2.781	1.064	2.363	1.640	14.053	3.614
couple	35	18.606	3.066	3.975	5.463	6.442	7.518
couple	36	5.000	4.842	5.000	6.138	1.114	8.064
couple	37	18.350	2.611	.925	1.290	13.581	1.997
couple	38	16.019	4.653	1.275	1.842	11.852	4.050
chat	39	3.850	3.397	.238	.846	21.473	3.620
chat	40	5.644	2.394	1.575	2.061	17.904	5.413
chat	41	15.294	5.140	1.150	2.820	24.297	9.339
chat	42	3.400	1.811	.325	.652	20.958	1.855
chat	43	18.419	4.448	1.163	3.671	25.125	14.674
chat	44	12.769	4.474	.125	.644	27.139	3.889
chat	45	16.525	5.798	1.288	2.404	24.425	10.090
chat	46	16.519	5.376	1.863	3.575	22.079	14.260
chat	47	14.256	5.173	2.738	4.078	16.873	10.983
chat	48	6.825	2.278	.488	.729	22.318	2.131
chat	49	6.475	2.544	.675	1.868	21.149	6.381
chat	50	14.975	6.314	1.638	1.989	22.117	7.492
chat	51	4.606	1.588	.538	.745	20.899	2.632
chat	52	16.588	4.950	1.063	1.503	25.762	6.908
chat	53	10.238	3.825	2.500	2.837	16.332	8.212
chat	54	7.638	2.940	3.488	4.704	10.635	11.153