Words versus actions as a means to influence cooperation in social dilemma situations

**Ganna Pogrebna · David H. Krantz ·
Christian Schade · Claudia Keser**

Abstract   We use a sequential voluntary contribution game to compare the relative impact of a first-mover's non-binding announcement versus binding commitment on cooperation. We find that a non-binding announcement and a binding commitment increase individual contributions to a similar extent. Since announced contributions systematically exceed commitments, in sessions with a non-binding announcement, second-movers tend to contribute more to the group activity than in sessions with a binding commitment. Yet, second-movers appear to be more motivated towards achieving a social optimum when the first-mover uses commitment. We also find that a non-binding announcement has a higher impact on individual propensity to cooperate than the ex post contribution of the first-mover. However, the failure to make announced contributions decreases cooperation even though the first-mover is reassigned in every period.

G. Pogrebna (✉)
Economics Department, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK
e-mail: G.Pogrebna@warwick.ac.uk

D. H. Krantz
Department of Psychology, Columbia University, 406 Schermerhorn Hall, New York, NY 10027, USA

C. Schade
Institute for Entrepreneurial Studies and Innovation Management, School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Strasse 1, 10178 Berlin, Germany

C. Keser
Faculty of Economic Sciences, Georg-August-Universität Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany

1 Introduction

When choosing whether to cooperate with others, people often face a variety of incentives and disincentives. The decision to cooperate may be influenced by many factors including prior commitments, preliminary communication, and even perceived social norms. The main purpose of this article is to compare and contrast two possible means to foster group cooperation: one-way binding commitment versus one-way non-binding announcement of the future commitment.

We conduct three treatments of a simple voluntary contribution game which has a unique Nash equilibrium in pure strategies (e.g., Isaac et al. 1984). In the first treatment, all players make simultaneous and binding decisions about how much money they want to contribute to the group activity without communicating with each other. In the second treatment, one of the players (henceforth, a leader) moves first and has an opportunity to influence cooperation by taking action (e.g., Moxnes and Van der Heijden 2003; Güth et al. 2007).[1] In this case, the leader makes a prior binding commitment (to contribute or not) and sets an example for the other group members (henceforth, followers) before they make their decisions.

In the third treatment, the leader also makes the first move. However, in contrast to the second treatment, she uses words rather than actions to influence cooperation. In this treatment, the leader has an opportunity to announce the level of her intended contribution to the followers, but is not bound by that announcement. By comparing the results from all three treatments, we explore the relative impact of a non-binding announcement versus a binding commitment on the level of individual and group contributions. We check whether participants are more likely to play the Nash equilibrium in a case when leaders use announcement rather than commitment. We also identify factors that influence individual decisions to contribute in different treatments using an econometric analysis which accounts for the unobserved heterogeneity of participants.

This article is related to two large streams of literature: (a) literature which studies how non-binding announcements which do not have a direct impact on players' payoffs (*cheap talk*) influence decision making and (b) literature on voluntary contribution games. Several papers investigate the impact of cheap talk in strategic games from a theoretical prospective. Particularly, Crawford and Sobel (1982) and Green and Stokey (1980) identify conditions when cheap talk leads to a transfer of valuable private information in a sender–receiver environment. Farrell (1987, 1988) analyzes the communicational intentions in games with complete information. He finds that while cheap talk does not necessarily guarantee reaching Nash equilibrium, it may serve as an efficient coordination tool. Rabin (1994) shows that in iterated play when players use cheap talk for a long time, in every equilibrium of every game each player

---

[1] In the economics literature, this treatment is often referred to as *leading by example*.

receives a payoff which is superior to her worst Pareto-efficient Nash equilibrium.[2] Baliga and Morris (2002) examine the effects of one-sided incomplete information on cheap talk and identify conditions when players coordinate on efficient Nash equilibria. In a recent study, Ottaviani and Sørensen (2006) explore the reputational considerations in a game with cheap talk. They find that relative reputation is an important factor which affects individual decisions.

From the empirical perspective, non-experimental studies of cheap talk are very rare due to the complexity of obtaining such data from the field. For example, Genesove and Mullin (2001) show that cheap talk fosters price collusion among companies involved in the sugar refining cartel. The impact of cheap talk has also been studied experimentally in a wide variety of games. Charness and Grosskopf (2004) find that cheap talk enforces coordination in a 2×2 stag hunt game. Blume and Ortmann (2007) show that it facilitates coordination on the Pareto-dominant equilibrium in games with many players. Sally (1995) and Balliet (2010) provide a detailed review of the literature on communication (including cheap talk) in social dilemmas. They show by means of meta-analysis that an overwhelming number of studies in this area find a robust positive effect of communication on the level of cooperation in groups.

Research on voluntary contribution games is a large and constantly growing field. To date, theoretical and experimental studies on voluntary contribution games have concentrated either on the effects of the binding commitment in the leader–follower setting or on the impact of a non-binding announcement. However, to our best knowledge, there have been no attempts to analyze the relative effects of the two. Particularly, in a theoretical paper, Hermalin (1998) develops a model where a leader uses a prior binding commitment to set an example for the rest of the group under conditions of asymmetric information. He shows that the leader's example fosters cooperation. When the followers observe that the leader exerts effort toward attaining the group goal, they tend to increase their effort levels. Several empirical studies, inspired by Hermalin's results, explore the impact of leaders' binding commitments in voluntary contribution experiments (e.g., Gächter and Renner 2003; Moxnes and Van der Heijden 2003; Arbak and Villeval 2007; Levati et al. 2007; Güth et al. 2007; Potters et al. 2007; Gächter et al. 2010). They provide robust evidence that leaders' positive prior commitments to the group activity increase followers' contributions and facilitate cooperation.

Bochet et al. (2006) explore the impact of different variations of non-binding announcements and communication in a voluntary contribution game when all group members move simultaneously. In their experiment, participants from the same group can engage in cheap talk: (a) by speaking face-to-face; (b) by verbal PC-to-PC chatting using the electronic chat room; and (c) by numerical communication via computer terminals. Bochet et al. (2006) show that face-to-face as well as verbal PC-to-PC communication increases efficiency in the voluntary contribution game to a similar extent. However, numerical communication does not have an impact on contributions.

This article contributes to the existing literature in several ways. First, our experimental design allows comparing two possible means of influencing cooperation:

---

[2] Farrell and Rabin (1996) and Crawford (1998) provide detailed overview of the cheap talk literature.

non-binding announcement versus binding commitment. Second, we extend the experimental literature on cheap talk and voluntary contribution games by analyzing whether and to what extent differences between treatments can be explained by the differences in characteristics of announcements and commitments. In addition, our analysis allows us to determine the effect of unobserved heterogeneity of individual participants in the voluntary contribution game on the level of their contributions.

We find that participants contribute significantly higher monetary amounts to the group activity in the treatment with a non-binding announcement and in the treatment with a binding commitment when compared with the treatment where all players move simultaneously. Furthermore, a non-binding announcement generates an increase in contributions which is at least as high as in the treatment with a binding commitment.[3] This result suggests that leader's words and actions have essentially the same impact on cooperation.

We also find that followers contribute higher amounts after observing a non-binding announcement rather than a binding commitment. This observation can be explained by the fact that non-binding announcements are systematically higher than binding commitments. Nevertheless, our results indicate that a binding commitment creates more incentives for experimental participants to concentrate on achieving a socially optimal outcome than a non-binding announcement.

Followers are more likely to contribute positive amounts to the group activity if the leader announces her intention to contribute a positive amount. Furthermore, an announcement appears to be a more important determinant of individual decisions than the amount of the leader's subsequent contribution. Nevertheless, followers take the reputation of leaders into account. Contributions decline if the leader in the previous period has failed to carry out an announced plan.

The remainder of the article is structured as follows. Section 2 provides the design and the theoretical predictions of the experiment and describes the experimental procedure. Results of the empirical analysis are reported in Section 3. Section 4 concludes by discussing results of our analysis.

2 The experiment

2.1 Experimental design and theoretical predictions

We consider a simple iterated voluntary contribution game (e.g., Isaac et al. 1984). A group of N players participate in the game during $t \in [1, T]$ periods. At the beginning of period t, each player $i \in \{1, \ldots, N\}$ receives an initial endowment k and has an opportunity to contribute $c_i^t \in \{0, k\}$ to the group activity. In other words, in every period, players can contribute either all of their initial endowment ($c_i^t = k$) or nothing ($c_i^t = 0$). The payoff of player i in period t is given by:

---

[3] In both the treatments individual participants and groups, on average, fail to play according to the predictions of the Nash equilibrium. Furthermore, even though we observe a gradual decline in contributions as the game progresses participants' behavior fails to converge to the Nash equilibrium in iterated play.

$$\pi_i^t = k - c_i^t + \frac{N-1}{N} \cdot \sum_{i=1}^{N} c_i^t \qquad (1)$$

This voluntary contribution game has one Nash equilibrium in pure strategies. Since $\frac{N-1}{N} < 1$, the dominant strategy for every player is to contribute $c_i^t = 0$. If in period t all players in the group contribute nothing, the payoff of each player in this period is equal to the initial endowment ($\pi_i^t = k \, \forall \, i$). However, this outcome is not socially efficient because it fails to maximize the sum of individual payoffs of the group members ($\sum_{t=1}^{N} \pi_i^t$). Social efficiency is reached only if all players contribute $c_i^t = k$, yielding $\pi_i^t = k(N-1) \forall i$.

In this experiment, we design three treatments: *BASELINE*, *ACTIONS*, and *WORDS*. In the *BASELINE* treatment, we conduct a simple iterated voluntary contribution game, described above. In this treatment, N players decide on the amount of their individual contribution $c_i^t \in \{0, k\}$ simultaneously, independently and without communicating with each other. We use the *BASELINE* treatment as a control treatment in our analysis.

Each period of the *ACTIONS* treatment consists of three stages. In stage $s = 0$, one player out of N is randomly assigned the role of *leader* (l) and the other $N - 1$ players are assigned the roles of *followers* ($f \neq l$). The difference between the roles is that the leader has an opportunity to make the first move in the game. Followers can move only after observing the move of the leader.

In stage $s = 1$, the leader makes a binding decision about the amount of contribution $c_l^t \in \{0, k\}$ to the group activity. In stage $s = 2$, followers observe the amount of the leader's contribution c_l^t and decide on the amount of their individual contributions $c_f^t \in \{0, k\}$ simultaneously, independently and without communicating with each other.

The theoretical prediction for the *ACTIONS* treatment coincides with the theoretical prediction for the *BASELINE* treatment. Assuming that a payoff maximization mechanism is common knowledge, since $\frac{N-1}{N} < 1$ in stage $s = 2$, the dominant strategy for each of $N - 1$ followers is to free-ride ($c_f^t = 0$). A rational payoff-maximizing leader anticipates this response and, therefore, contributes $c_l^t = 0$ in stage $s = 1$. Similarly to the *BASELINE* treatment, a socially efficient outcome is reached in the *ACTIONS* treatment when all players contribute k to the group activity.

In the *WORDS* treatment, each period incorporates three stages. Similar to the *ACTIONS* treatment, in stage $s = 0$, each of the N players in the group is assigned the role of leader (l) or follower (f), $f \neq l$. In this treatment, the leader is also the first mover. However, instead of making a binding contribution to the group activity, the leader has an opportunity to make a non-binding announcement of her future contribution.

In stage $s = 1$, the leader makes an announcement $m \in \{0, 1\}$ to the $N - 1$ followers. If $m = 0$, the leader announces her plan to contribute $c_l^t = 0$ to the group activity, if $m = 1$ the leader announces her plan to contribute $c_l^t = k$. This non-binding announcement is communicated to all $N - 1$ followers in the group. In stage $s = 2$, all group members (including the leader) make binding decisions about the amount of their contributions simultaneously, independently and without communicating with each other.

Since the leader's announcement in the *WORDS* treatment is non-binding, the one-way pre-game communication is essentially cheap talk which does not alter the prediction of the Nash equilibrium in pure strategies. In other words, all group members should contribute nothing to the group activity. Similar to the *BASELINE* and the *ACTIONS* treatments, a socially efficient outcome in the *WORDS* treatment is a situation when all members of the group contribute k.

The majority of voluntary contribution experiments allow any fraction of initial endowment to be contributed to the group activity (e.g., Isaac et al. 1984; Palfrey and Prisbrey 1997; Bochet et al. 2006). In this experiment, we concentrate on a situation when players face a binary choice between contributing all of their initial endowment or nothing for three main reasons described below.

First, all-or-nothing design allows us to investigate the relative impact of the leader's binding and non-binding first move on contributions in an extreme case when contributing means giving up the entire endowment. Previous experimental studies (e.g., Isaac et al. 1984; Palfrey and Prisbrey 1997; Moxnes and Van der Heijden 2003; Bochet et al. 2006; Güth et al. 2007) provide robust evidence of positive contributions in voluntary contribution games. One of the possible explanations for this behavior is that experimental participants make errors (e.g., Offerman et al. 1998). Intuitively, such errors are more likely to occur in the case when any fraction of initial endowment can be contributed to the group activity when compared with the all-or-nothing case. To minimize the possibility of stochastic errors, we limit participants' action space to the binary choice decisions.

Second, in the *WORDS* treatment, we are interested in analyzing a situation when unfulfilled announcements may have serious consequences for a leader's reputation. For example, consider a case when the leader announces a future plan to contribute k but has an opportunity to contribute $\frac{k}{2}$. Intuitively, the contribution of $\frac{k}{2}$ after announcing k should damage the leader's reputation to a lower extent than the contribution of 0. Finally, the binary nature of each player's action space simplifies the decision problem for the participants.

2.2 Experimental procedure

We have conducted six sessions of the experiment (two sessions per treatment). Twelve participants took part in each session, yielding a total of 72 participants. We have used between-subject design. In each session of the experiment, participants took part in only one treatment of the voluntary contribution game (*BASELINE*, *ACTIONS*, or *WORDS*).

Participants were recruited via the online invitation system at Humboldt-Universität zu Berlin. The majority of participants were students at Humboldt-Universität zu Berlin. Less than 1/5 of participants (18.1%) studied either Economics or Business Administration and had previous exposure to game theory.

The sample was relatively balanced in terms of gender composition. 54.2% of participants were female and 45.8% were male. The average age of participants was 26 years with a median of 25 and a standard deviation of 6 years. 73.6% of participants had previous experience with decision making experiments. However, none of them

had played a voluntary contribution game before. The majority of participants (86.1%) reported an annual income below € 15,000.

All experimental sessions were conducted at the experimental laboratory of Humboldt-Universität zu Berlin. The experiment was computerized using the z-Tree software (Fischbacher 2007). Upon their arrival at the experimental laboratory, participants were seated at individual workstations equipped with a personal computer, scratch paper, and a pen. The workstation of each participant was separate and could not be seen by other participants and/or the experimenter.

The experiment consisted of two experimental tasks and a post-experimental questionnaire. In the first experimental task, participants were subjected to the iterated voluntary contribution game. In the second task, they took part in the Holt and Laury (2002) risk attitude elicitation procedure. Participants received hard copies of experimental instructions for each task separately.[4]

Instructions were read aloud by the experimenter. After listening to the experimenter, participants were given time to study the instructions individually and ask questions, which were answered privately. Any communication among participants during the experiment was prohibited.

Irrespective of the treatment, at the beginning of the first task, participants were randomly divided into groups of three people each ($N = 3$). In every session, participants played 20 periods of the voluntary contribution game ($t \in [1, 20]$). Participants were not informed about the exact number of rounds that they were about to play. However, they were informed that the first experimental task will not take more than 30 min. In other words, while participants did not know the exact number of rounds, they knew that they will be playing a finite and a relatively short game which could end in any period. Group compositions remained constant for the duration of the first experimental task. In the *ACTIONS* and the *WORDS* treatment roles were reassigned in every period using a random draw. At the beginning of each period, all the participants received an initial endowment of € 10 ($k = 10$). At the end of the experiment, participants received the payoff from one randomly chosen period of the voluntary contribution game.

In all three treatments of the experiment, participants received full feedback about the outcome of their decisions at the end of every round of the game. Particularly, in the *BASELINE* treatment they received information about (a) their own individual contributions, (b) individual decisions of other players in their group; (c) sum of all contributions in the group; (d) their individual payoffs in the round. To preserve confidentiality, at the beginning of the first experimental task, every player in the group was randomly assigned an ID (A, B, or C) by the computer program. During the voluntary contribution game, players were identified only by their IDs.[5]

In the *ACTIONS* treatment, participants received feedback about (a) their individual contributions; (b) ID of the leader; (c) individual contributions of other players

---

[4] Experimental instructions are provided in the Electronic supplementary material.

[5] For example, at the end of each period, Player A received information about individual contributions of Players B and C. However, she did not know the names or any other personal information about players in her group such as their seat numbers and etc. Player B (C) received the same feedback about individual contributions of Players A and C (B).

in their group (d) sum of all contributions in their group; (e) their individual payoffs in the round. In the *WORDS* treatment, players received information about (a) their individual contributions, (b) ID of the leader; (c) the announcement of the leader; (d) individual contributions of players in their group (including the leader); (e) sum of all contributions in their group; (f) their individual payoffs in the round. We used neutral language to identify leaders and followers, i.e., followers were labeled as TYPE 1 players and leaders as TYPE 2 players. To explore the impact of the reputation of individual leaders, types were assigned at random by the computer program and reported to all players at the beginning of each period. In other words, in the *ACTIONS* and the *WORDS* treatments, followers knew the ID of the leader before they made their decisions.

To avoid wealth effects, the payoff from both experimental tasks was determined at the end of the experiment. Upon completion of the experimental tasks, participants received a questionnaire with demographic questions. The whole experimental procedure, including the questionnaire, lasted approximately 1 h. Average earnings of the participants were € 18.50 with a median of € 18.30 and a standard deviation of € 4.89.[6]

3 Results

In this section, we explore the relative impact of a binding commitment versus a non-binding announcement on the level of cooperation in the voluntary contribution game. First, we check whether experimental participants across all the three treatments of our experiment behave according to the predictions of the Nash equilibrium. We also analyze whether and to what extent individual and group contributions in *BASELINE*, *ACTIONS*, and *WORDS* change in the iterated play. Second, we compare and contrast decisions of leaders and followers in the *ACTIONS* and the *WORDS* treatment and try to explain the observed differences. Third, we provide a detailed analysis of the leader's announcements and followers' responses in the *WORDS* treatment. Finally, we identify the main determinants of individual contributions by conducting an econometric analysis of the data.

3.1 Treatment effects

This section is devoted to the analysis of treatment effects. We explore relative differences in contributions across all the three treatments in our experiment.

Similar to the findings in the previous literature on voluntary contribution games (e.g., Moxnes and Van der Heijden 2003; Güth et al. 2007), our data suggest that participants at both individual and group levels do not behave according to the predictions of the Nash equilibrium in pure strategies. According to Fig. 1, contributions in *BASE-*

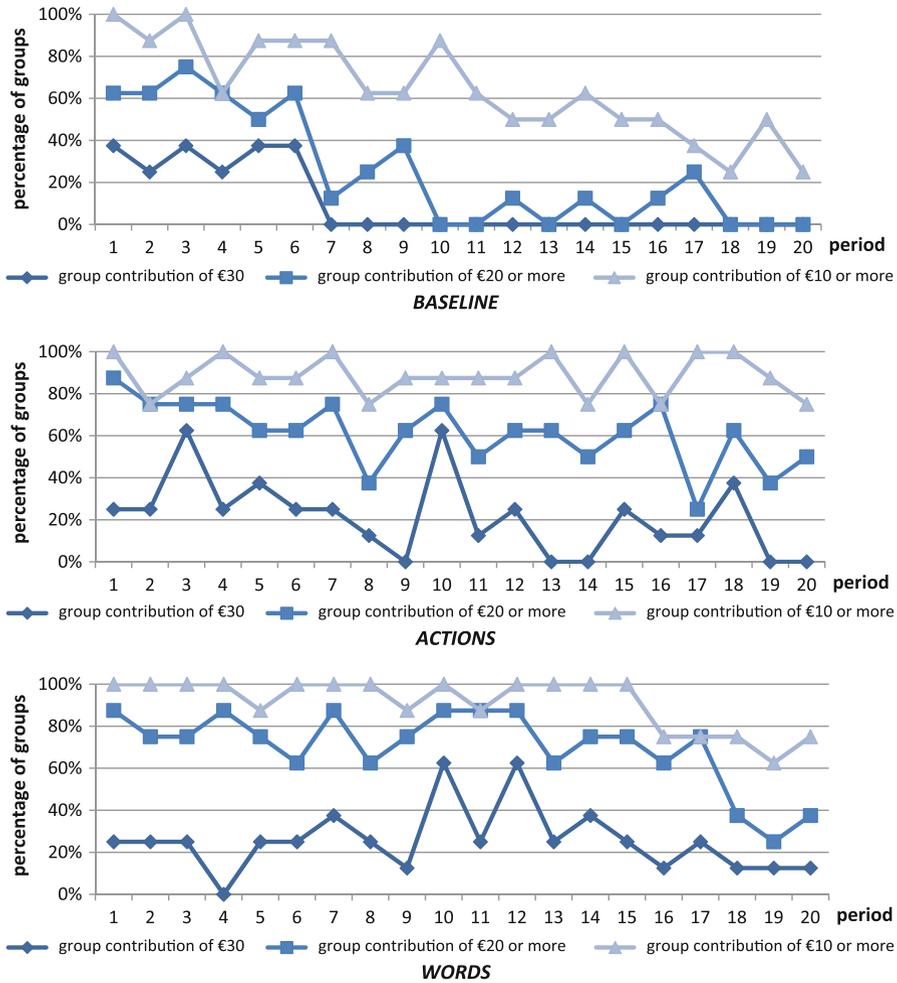[6] At the time of the experiment, the exchange rate was € 1=$1.56.

Fig. 1 Observed positive group contributions in iterated play.[8]

LINE, ACTIONS, and *WORDS* decline with iterated play.[7] However, participants in all the treatments tend to contribute positive amounts to the group activity in all rounds

---

[7] Decline in the level of contributions is a robust finding in experimental literature. Economic research offers two main explanations for this phenomenon. One argues that strategies of players are well-defined from the beginning of the iterated play. However, these strategies are state-contingent and, therefore, depend on the history of play. The other explanation maintains that experimental participants learn to play the game which makes them gradually refine their strategies. Current experimental evidence is mixed: some papers provide support for the state-contingent hypothesis, while others support the learning hypothesis. Andreoni and Croson (2008) and Muller et al. (2008) provide a comprehensive review of this literature.

[8] Group contribution of 0 or more is not shown since all observed group contributions fall under this category
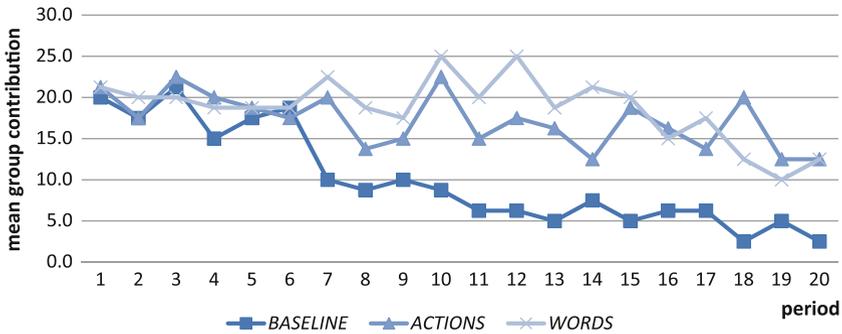
Fig. 2 Mean group contributions in *BASELINE*, *ACTIONS*, and *WORDS* treatments

of the game. Interestingly, while in the later rounds of the *BASELINE* treatment more and more participants appear to switch to contributing nothing to the group activity, there appears to be no convergence to the equilibrium prediction in the *ACTIONS* and the *WORDS* treatment.

Figure 2 shows the dynamics of mean group contributions across 20 periods.[9] Apparently, mean contributions are similar across all treatments during the first six periods of the game. After that, contributions in the *ACTIONS* and the *WORDS* treatment decline at a lower rate than contributions in the *BASELINE* treatment. Furthermore, in periods 7–14, contributions in the *WORDS* treatment tend to be higher than in the other two treatments. However, in the last six periods (periods 15–20), participants contribute similar amounts in the *ACTIONS* and the *WORDS* treatment. Yet, the contributions in *ACTIONS* and *WORDS* appear to be much higher than in *BASELINE*.

Result 1 Non-binding announcement increases the level of contributions to the group activity in the iterated play.

Table 1 depicts mean individual and group contributions in each of the three treatments of the experiment. According to Table 1, mean individual and group contributions across all 20 periods are highest in the *WORDS* treatment and lowest in the *BASELINE* treatment. Contributions in the *ACTIONS* treatment are higher than in the *BASELINE* treatment but lower than in the *WORDS* treatment. Results of the non-parametric Cuzick trend test across all three treatments show that there is an upward *BASELINE–ACTIONS–WORDS* trend in the individual and group contributions data (Cuzick test $p < 0.001$ for both the individual and the group comparison).

Mean contributions appear to be strikingly similar in all three treatments at the beginning of the game and in *ACTIONS* and *WORDS* at the end of the game. Therefore, we divide the data into three parts: (a) data from periods 1–6, (b) data from periods 7–14, and (c) data from periods 15–20 (see Table 1). Re-applying the trend test procedure to each part of the data provides further evidence for the upward *BASE-LINE–ACTIONS–WORDS* trend in periods 7–14 (Cuzick test $p < 0.001$ for both individuals and groups) but not in periods 1–6 (Cuzick test $p > 0.44$ for individ-

---

[9] Individual contributions follow the same pattern.

Table 1 Contributions in *BASELINE, ACTIONS*, and *WORDS*[a]

Time	Players	Treatment	Mean individual contribution (standard error)	Standard deviation of individual contributions	Mean group contribution (standard error)	Standard deviation of group contributions
Periods 1–6	All	*BASELINE*	6.11 (0.41)	4.89	18.33 (1.50)	10.38
		ACTIONS	6.53 (0.40)	4.78	19.58 (1.40)	9.67
		WORDS	6.53 (0.40)	4.78	19.58 (1.03)	7.13
	Leaders	*ACTIONS*	5.63(0.72)	5.01	–	–
		WORDS	3.13 (0.68)	4.68	–	–
	Followers	*ACTIONS*	6.98 (0.47)	4.62	–	–
		WORDS	8.23 (0.39)	3.84	–	–
Periods 7–14	All	*BASELINE*	2.60 (0.32)	4.40	7.81(0.82)	6.54
		ACTIONS	5.47 (0.36)	4.99	16.56 (1.12)	8.95
		WORDS	7.03 (0.33)	4.58	21.10 (1.02)	8.19
	Leaders	*ACTIONS*	5.31(0.63)	5.03	–	–
		WORDS	5.47 (0.63)	5.02	–	–
	Followers	*ACTIONS*	5.55 (0.44)	4.99	–	–
		WORDS	7.81 (0.37)	4.15	–	–
Periods 15–20	All	*BASELINE*	1.52(0.30)	3.61	4.58(0.89)	6.17
		ACTIONS	5.21 (0.42)	5.01	15.63 (1.26)	8.73
		WORDS	4.86 (0.42)	5.02	14.58 (1.49)	10.31
	Leaders	*ACTIONS*	4.38(0.72)	5.01	–	–
		WORDS	3.75 (0.71)	4.89	–	–
	Followers	*ACTIONS*	5.63 (0.51)	4.99	–	–
		WORDS	5.42 (0.51)	5.01	–	–
Periods 1–20	All	*BASELINE*	3.33 (0.22)	4.72	10.00(0.76)	9.58
		ACTIONS	5.71 (0.23)	4.95	17.19 (0.73)	9.19
		WORDS	6.23 (0.22)	4.85	18.69 (0.71)	8.98
	Leaders	*ACTIONS*	5.13(0.40)	5.01	–	–
		WORDS	4.25 (0.39)	4.96	–	–
	Followers	*ACTIONS*	6.00 (0.27)	4.91	–	–
		WORDS	7.21 (0.25)	4.49	–	–

[a] All contributions are in euros

uals and $p > 0.75$ for groups). In periods 15–20, Cuzick test suggests the upward trend between *BASELINE* and *ACTIONS* which has a similar ranking with *WORDS* ($p < 0.001$ for both individuals and groups). This result suggests that differences in participants' behavior across treatments develop as the game progresses. This means that participants do not react to the treatment variation immediately but rather adjust to different decision contexts over time.

Result 2 Non-binding announcement and binding commitment increase contributions to a similar extent.

According to Table 1, on average, all experimental participants in the *WORDS* treatment contribute almost twice as much money to the group activity as in the *BASELINE* treatment across all periods of the game. Results of the Wilcoxon–Mann–Whitney test (Mann and Whitney 1947) suggest that contributions in periods 1–20 are statistically significantly higher in the *WORDS* treatment than in the *BASELINE* treatment ($p < 0.001$ for both the individual and the group level). This difference in participants' behavior is primarily observed in later periods of the game. Table 1 provides summary statistics of mean individual and group contributions in periods 1–6, periods 7–14, and periods 15–20 for all the treatments. In periods 1–6, contributions in *BASELINE* and *WORDS* are very similar (Wilcoxon–Mann–Whitney test $p > 0.45$ for individuals and $p > 0.71$ for groups). However, in periods 7–14 and 15–20, contributions in the *WORDS* treatment become statistically significantly higher (all probabilities in a series of Wilcoxon–Mann–Whitney tests are less than 0.001 for individuals and groups).

We also find that across all 20 periods of the game in the *ACTIONS* treatment participants contribute statistically significantly more than in the *BASELINE* treatment. Results of the Wilcoxon–Mann–Whitney test indicate that mean individual ($p < 0.001$) and group ($p < 0.001$) contributions are statistically significantly lower in the *BASELINE* treatment than in the *ACTIONS* treatment. This finding is consistent with the previous literature (e.g., Moxnes and Van der Heijden 2003; Güth et al. 2007). Similar to the *WORDS* treatment (see Table 1), contributions in the *ACTIONS* treatment are similar to contributions in the *BASELINE* treatment in periods 1–6 (Wilcoxon–Mann–Whitney test $p > 0.45$ for individuals and $p > 0.56$ for groups). However, in periods 7–14 and 15–20, *ACTIONS* contributions are statistically significantly higher (all probabilities in a series of Wilcoxon–Mann–Whitney tests are less than 0.001 for individuals and groups).

Interestingly, while in periods 7–20, we observe a significant cross-treatment effect between the *BASELINE* and the *ACTIONS* treatment and between the *BASELINE* and the *WORDS* treatment, the comparison between *ACTIONS* and *WORDS* yields a different result. In periods 7–14, the level of contributions in the *WORDS* treatment is statistically significantly higher than in the *ACTIONS* treatment (Wilcoxon–Mann–Whitney test $p < 0.002$ for individuals and $p < 0.005$ for groups). However, in periods 15–20, contributions in the *ACTIONS* and the *WORDS* treatment are essentially the same (Wilcoxon–Mann–Whitney test $p > 0.54$ for individuals and $p > 0.65$ for groups). The similarity between the *ACTIONS* and the *WORDS* treatment in the last six periods of the game is apparent in Fig. 2.

This result suggests that an increase in the level of contributions in a social dilemma situation can be reached by the leader's non-binding announcement and does not require a binding commitment. Furthermore, in the iterated play, a non-binding announcement might be sufficient for reaching a higher level of cooperation than a binding commitment. Nevertheless, despite the different dynamics of a non-binding announcement and a binding commitment, they both eventually lead to the same level of contributions.

3.2 Comparative analysis of leaders' and followers' behavior

In this section, we consider experimental treatments where leaders make the first move in the game (i.e., the *ACTIONS* and the *WORDS* treatment). We investigate whether and to what extent the role that participant i plays has an impact on her decisions. Particularly, we conduct a comparative analysis of the level of individual contributions and individual payoffs of leaders and followers.

Result 3 Leaders contribute similar amounts in treatment with non-binding announcement and binding commitment. However, followers make higher contributions when leaders make non-binding announcements.

Table 1 provides information about the mean leaders' and followers' contributions in the *ACTIONS* and the *WORDS* treatment. The Wilcoxon–Mann–Whitney test conducted on the individual data shows that contributions of leaders are essentially the same in the *ACTIONS* and the *WORDS* treatment ($p > 0.10$) across all periods of the game. Interestingly, in periods 1–6 leaders in the *ACTIONS* treatment contribute statistically significantly more than leaders in the *WORDS* treatment (Wilcoxon–Mann–Whitney test $p < 0.02$). However, later in the game, leaders in both the treatments contribute similar amounts to the group activity (Wilcoxon–Mann–Whitney test $p > 0.84$ in periods 7–14 $p > 0.52$ in periods 15–20).

Despite the similarity in the leaders' behavior, followers in the *WORDS* treatment show a systematically higher level of contributions than followers in the *ACTIONS* treatment for the majority of periods (see Table 1). Wilcoxon–Mann–Whitney test results show that followers contribute statistically significantly higher monetary amounts in the *WORDS* treatment than in the *ACTIONS* treatment across all 20 periods ($p < 0.002$). However, the robustness check reveals an interesting pattern in the data. At the beginning of the experiment, followers make higher contributions in *WORDS* than in *ACTIONS* (Wilcoxon–Mann–Whitney test $p < 0.05$ in periods 1–6 and $p < 0.001$ in periods 7–14). However, in the last six periods (periods 15–20) followers in both treatments contribute similar amounts (Wilcoxon–Mann–Whitney test $p > 0.76$).

Since followers in *WORDS* and *ACTIONS* receive qualitatively different information before making their decisions, we can formulate the following hypothesis. If the observed difference in the followers' contributions across all 20 periods is inspired by the treatment variation, it must be the case that the overall level of leaders' non-binding announcements in the *WORDS* treatment systematically exceeds the level of leaders' binding commitments in the *ACTIONS* treatment.[10] We can also expect the differences between non-binding announcements and binding commitments to be particularly profound in periods 1–14 but then fading away in periods 15–20.

We check this hypothesis by conducting a non-parametric comparison of leaders' non-binding announcements versus binding commitments across two treatments. We find that while on average leaders in the *WORDS* treatment announce a plan to

---

[10] This hypothesis was suggested to us by an anonymous referee.

contribute € 8.06, leaders in the *ACTIONS* treatment make a binding first contribution of € 5.13 across all 20 periods of the game. Results of Wilcoxon–Mann–Whitney test conducted on individual data from all periods suggest that the difference between announcements and commitments is statistically significant ($p < 0.001$).

This result remains the same after we apply a robustness check by running separate tests for all parts of the data. In the *ACTIONS* treatment, leaders on average contribute € 5.63 in periods 1–6, € 5.31 in periods 7–14, and € 4.38 in periods 15–20. In the *WORDS* treatment, leaders on average announce a plan to contribute € 7.71 in periods 1–6, € 8.59 in periods 7–14 and € 7.71 in periods 15–20. Results of Wilcoxon–Mann–Whitney test suggest that the level of non-binding announcements is statistically significantly higher than the level of binding commitments in all three parts of the data: $p < 0.04$ in periods 1–6; $p < 0.001$ in periods 7–14 and $p < 0.001$ in periods 15–20. Therefore, our hypothesis is partially confirmed. We observe significant differences between announcements and commitments across all periods and at the beginning of the game. However, despite our expectation, the differences remain in the last six periods. This can be explained by analyzing behavior of leaders in the *WORDS* treatment. Particularly, it might be the case that even though the leaders make high announcements throughout the game, in the later periods, they deviate from their announcements and contribute zero to the group activity more frequently. This damages their reputation and followers decrease their contributions despite leaders' high announcements. In Sect. 3.3, we conduct a detailed analysis of leaders' and followers' behavior in the *WORDS* treatment and check whether this explanation is correct.

Result 4 A non-binding announcement has a positive effect on leaders' payoffs but does not change followers' payoffs compared with a binding commitment.

The analysis of leaders' and followers payoffs indicates that followers in both the *ACTIONS* (mean payoff is equal to € 15.45) and the *WORDS* treatment (mean payoff is equal to € 15.23) receive similar payoffs. The results of the Wilcoxon–Mann–Whitney test conducted on individual data shows that there is no statistically significant cross-treatment effect on the payoff of the followers ($p < 0.50$). However, pre-game communication significantly improves the earnings of leaders ($p < 0.001$). They receive, on average, almost € 2.00 more in the *WORDS* treatment (mean payoff is equal to € 18.21) than in the *ACTIONS* treatment (mean payoff is equal to € 16.33).

Result 5 Leaders contribute less than followers in treatment with non-binding announcement, while contributions of leaders and followers are very similar in treatment with binding commitment.

Recall from Sect. 2 that in both the *ACTIONS* and the *WORDS* treatment leaders are randomly re-assigned in every period by the computer program. This means that each participant plays the voluntary contribution game both as a leader and as a follower several times. According to Table 1, on average leaders contribute € 5.13 and followers € 6.00 across all 20 periods in the *ACTIONS* treatment. Non-parametric analysis conducted on the mean individual contributions

suggests that leaders' and followers' contributions in the *ACTIONS* treatment are not statistically significantly different (Wilcoxon signed-rank test $p < 0.51$).[11] This result remains the same in all parts of the data: in periods 1–6, 7–14, and 15–20.

Table 1 also shows that in the *WORDS* treatment leaders on average contribute € 4.25 to the group activity, while followers' mean contribution is equal to € 7.21 across all the periods. The difference between individual mean contributions is statistically significant (Wilcoxon signed-rank test $p < 0.001$). This means that, in the *WORDS* treatment, followers contribute more money to the group activity than leaders. This result is the same in all three parts of the data: in periods 1–6, 7–14, and 15–20.

This finding is particularly interesting because the same participants played the roles of leaders and followers. It may have two possible explanations. First, a non-binding announcement catalyzes the impact of imposed identity, which induces different patterns of behaviors when participants play as leaders and as followers. Second, in the *WORDS* treatment, leaders take advantage of their first move. Particularly, they announce plans to contribute to the group activity and then free-ride exploiting the fact that the followers observe the leader's actual decision only ex post. We elaborate on these explanations below and in Sect. 3.3.

Result 6 Leaders earn more than followers in the treatment with non-binding announcement, while leaders and followers receive similar payoffs in the treatment with binding commitment.

According to the results of the Wilcoxon signed-rank test (Wilcoxon 1945), followers fail to take advantage of a leaders' first move in the *ACTIONS* treatment. The difference between the payoffs of followers (mean payoff is equal to € 15.45) and leaders (mean payoff is equal to € 16.33) in this treatment is not statistically significant ($p < 0.08$).[12] However, in the *WORDS* treatment leaders (mean payoff is equal to € 18.21) receive, on average, € 3.00 more than followers (mean payoff is equal to € 15.23). Wilcoxon signed-rank test shows that this difference is highly statistically significant ($p < 0.001$). This finding suggests that leaders take advantage of the opportunity to make a non-binding announcement to the followers.

3.3 Non-binding announcements

In this section, we analyze whether and how leaders have used the opportunity to announce their contribution plans to the rest of the group in the *WORDS* treatment. In our data set, leaders have made announcements inconsistent with their ex post con-

---

[11] This result appears to be at odds with the findings in the previous literature (e.g., Güth et al. 2007) that leaders tend to contribute more than followers when leaders make the first binding commitment. This deviation from the previous literature can be explained by the fact that our design allows players to contribute only either their entire initial endowment or nothing ($c_i^t \in \{0, k\}$); whereas, for example, in Güth et al. (2007) participants may contribute any intermediate amount ($c_i^t \in \{0, k\}$).

[12] This result is different from the findings in the previous literature (e.g., Güth et al. 2007) and may result from the binary structure of contribution decisions in our experimental design (see Footnote 10).

tributions in 63 out of 160 cases (in 39.4% of rounds). This fraction is statistically significantly lower than 50% (binomial test $p < 0.01$) suggesting that leaders try to avoid appearing inconsistent.

In 62 cases, leaders have announced positive contributions of € 10 ($m = 1$) but have subsequently contributed nothing. There is only one case when a leader has informed the group of her plan to contribute 0 ($m = 0$) but has changed her mind and has contributed € 10. In order to explore the impact of non-binding announcements in iterated play, we formulate and test the following two hypotheses about leaders' behavior.

Our first hypothesis is that in iterated play, a rational leader should be exactly indifferent between making an announcement which is consistent and inconsistent with her ex post contribution. Recall from the experimental procedure described in Sect. 2 that group members are informed about the ID of the leader at the beginning of each period. In order to determine the content of her announcements in the iterated play, a rational leader will apply backward induction.[13] In the last period of the game ($t = T$), followers will free-ride irrespective of the announcement content. Therefore, in period $t = T$, leader's reputation will not have an impact on the followers' decisions. In period $t = T - 1$, a leader does not have an incentive to make an announcement consistent with her subsequent contribution because preserving a good reputation will not alter the outcome of the game in period $t = T$. Proceeding further by backward induction yields a theoretical prediction that rational leaders should treat a non-binding announcement as cheap talk. In other words, these leaders should be exactly indifferent between announcing plans consistent and inconsistent with their subsequent contributions.

This theoretical prediction has two possible interpretations. On the one hand, a strict interpretation suggests that a rational leader should randomize between announcing consistent and inconsistent plans with equal probability. In this case, we should observe a uniform distribution of frequencies of inconsistent announcements across all 24 leaders. In other words, each participant in the *WORDS* treatment should resort to inconsistent announcements in 50% of periods when this participant plays the role of the leader. On the other hand, a broad interpretation maintains that indifference may mean that leaders randomize between consistent and inconsistent announcements with any probability. In this case, any observed announcement strategy is possible.

Figure 3 summarizes frequencies of inconsistent announcements in the *WORDS* treatment across all leaders. It shows that there is substantial heterogeneity in leaders' propensities to make inconsistent announcements. It is apparent that while some leaders always make inconsistent announcements (frequency is equal to 1), others leaders are always consistent (frequency is equal to 0). The results of the Kruskal–Wallis equality of populations rank test with 23 degrees of freedom (Kruskal and Wallis 1952) confirm that leaders apply different tactics when deciding on whether or not to make an announcement consistent with their subsequent contribution ($p < 0.02$).

---

[13] Even though players are not informed about the exact number of periods in the game, they know that they are going to play a finite game which consists of several periods. We can apply the backward induction principle because each player should expect to play at least two periods of the game.
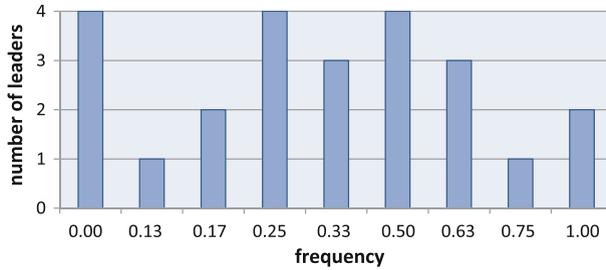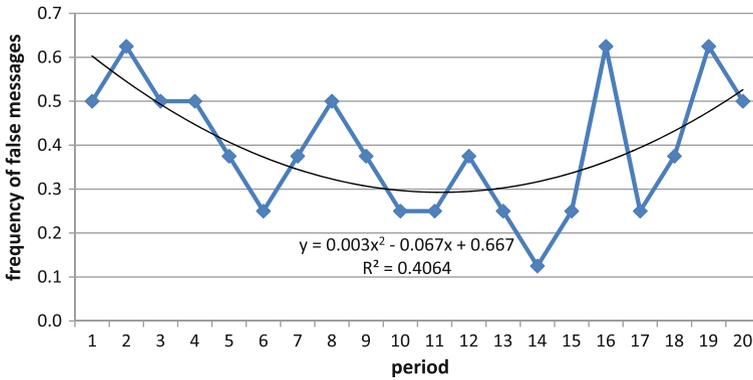
Fig. 3 Frequency of announcements inconsistent with subsequent contributions across leaders in the *WORDS* treatment

Only four participants (16.7%) have announced a different amount than their actual contribution in 50% of all periods when they played the role of the leader, 14 leaders (58.3%) contributed a different amount than announced less frequently and 6 (25.0%) more frequently. In other words, even when we look at the entire population of leaders they do not seem to randomize between consistent and inconsistent announcements with equal probability. This, however, may be due to the fact that not all leaders are rational.

A closer look at the data reveals that only 6 leaders in the population (25.0%) behave according to the prediction of Nash equilibrium derived in Sect. 2. In all periods of the game, they contribute nothing to the group activity. When we have limited our analysis to these six presumably rational leaders, it turned out that 1 of them made inconsistent announcements in 50% of cases, 2 in 62.5% of cases, 1 in 75% of cases, and 2 all the time. In other words, the majority of rational leaders have made inconsistent announcements in more than 50% of cases. Therefore, we can reject the strict interpretation of our first hypothesis.

Our second hypothesis is that if leaders have reputational concerns and believe that a positive reputation may increase their payoff, the frequency of inconsistent announcements should be relatively low at the beginning of the game but should increase as the game progresses. Figure 3 shows that the majority of leaders make announcements inconsistent with their subsequent actions at least once. This suggests that leaders might not treat non-binding announcements as cheap talk. Even in a situation when the game has a unique Nash equilibrium in pure strategies, there is no reason to believe that this equilibrium will be reached without some coordination between players. Under these circumstances, leaders might use a different rationale in making their announcement decisions. The leader might think that followers' rationalizable behavior in a subgame depends on their rationalizable beliefs, which can be influenced by the non-binding announcement.

If one allows for a possibility that leaders have such beliefs about followers' behavior, these beliefs may come into play in a variety of ways. For example, reputation considerations may become important. If leaders are concerned about their reputation, they will try to make announcements consistent with the subsequent contributions at the beginning of the game and then gradually divert to inconsistent announcements as the game progresses to take advantage of their first move.

Fig. 4 Frequency of announcements inconsistent with subsequent contributions by period

The analysis of the number of inconsistent messages in iterated play (see Fig. 4) reveals an interesting pattern. We observe that the frequency of inconsistent announcements has a U-shaped distribution. Since the number of messages inconsistent with leaders' subsequent decisions increase in the second half of the iterated play, our second hypothesis is partially confirmed.

This result may have two possible explanations. First, since leaders do not know the exact number of periods, they update their expectations about the length of the game as it progresses. At the beginning of the game they may not anticipate to be the leader more than once and start by sending inconsistent messages. However, as the game progresses they update their beliefs about the length of the game and realize the damage to their reputation. That is why they try to rehabilitate their image in the middle of the game by making consistent announcements. Nevertheless, towards the end of the game leaders increase the frequency of inconsistent announcements to take advantage of the pre-game communication. Another explanation might be that leaders behave irrationally when they decide on whether or not to announce to the rest of the group what they are actually going to do.

3.4 The determinants of individual contributions

In this section, we identify the determinants of individual contributions in all treatments of the experiment. Since the decision variable $c_i^t \in \{0, k\}$ is binary, we use a random intercept logit regression (e.g., Longford 1994) to explore factors that influence individual decisions. The dependent variable is a dummy variable y_i^t, specified as follows:

$$y_i^t = \begin{cases} 1, & \text{if } c_i^t = k \\ 0, & \text{if } c_i^t = 0 \end{cases} \qquad (2)$$

The probability that an individual i opts for contributing $c_i^t = k$ in period $t \in [1, T]$ is given by:

$$p\left(y_i^t = 1\right) = \frac{\exp\left(\beta_1 X1_i^t + \beta_2 X2_i^t + \cdots + \beta_M XM_i^t + \alpha_i\right)}{1 + \exp\left(\beta_1 X1_i^t + \beta_2 X2_i^t + \cdots + \beta_M XM_i^t + \alpha_i\right)}, \quad (3)$$

where $X1_i^t \ldots XM_i^t$ are explanatory variables described in Table 2; $\beta_1 \ldots \beta_M$ are regression coefficients and α_i is a vector capturing unobserved individual heterogeneity. The conditional log-likelihood function of the random intercept logit regression has the following form:

$$\mathrm{LL} = \prod_{i=1}^{N} \int_{-\infty}^{+\infty} \prod_{t=1}^{T} \left(\frac{\exp\left(\beta_1 X1_i^t + \beta_2 X2_i^t + \cdots + \beta_M XM_i^t + \alpha_i\right)}{1 + \exp\left(\beta_1 X1_i^t + \beta_2 X2_i^t + \cdots + \beta_M XM_i^t + \alpha_i\right)} \right) f(a)\, \mathrm{d}a \quad (4)$$

The log-likelihood function (4) is approximated using the adaptive quadrature method (Rabe-Hesketh et al. 2002).[14] Results of the random intercept logit regressions estimated with different number of explanatory variables are reported in Table 3.

According to Table 3, one variable influences participants' decisions in the *BASELINE* treatment. The propensity to make a positive contribution is higher in the early periods of the game (variable **PERIOD**). In other words, in the *BASELINE* treatment, participants take into account incentive consequences of playing in a certain period of the game when deciding on the amount of their contributions. They are more likely to contribute $c_i^t = k$ in the early periods and switch to $c_i^t = 0$ toward the end of the game even though they do not receive information about the exact number of periods.

Table 3 reports that in the *ACTIONS* treatment, the propensity to make a positive contribution decreases as the game progresses (variable **PERIOD**). Followers are more likely to contribute their entire initial endowment than leaders (variable **ROLE**). Most importantly, the leader's binding contribution amount has a highly statistically significant effect on the individual contributions in the group (variable **LCONTR**). Particularly, the higher the contribution of the leader, the more likely group members are to contribute $c_i^t = k$.

According to Table 3, five explanatory variables have a significant impact on participants' contributions in the *WORDS* treatment. Participants in the *WORDS* treatment appear to take more factors into account than in the *BASELINE* and the *ACTIONS* treatment. This suggests that they face a more complex decision problem in the *WORDS* treatment than in the other two treatments. Similar to the *BASELINE* and the *ACTIONS* treatment, participants are more likely to contribute their entire initial endowment to the group activity in early periods of the game (variable **PERIOD**) in the *WORDS* treatment. Furthermore, like in the *ACTIONS* treatment, followers are more likely to contribute than leaders (variable **ROLE**).

---

[14] The estimation has been conducted using the GLLAMM plug-in for the Stata 10.0 package. In addition to the two-level model with unobserved individual heterogeneity, specified above, we have estimated a two-level model with a random intercept at the level of a group in all treatments. We have also estimated three-level models with random intercepts at the level of individual participants and their respective roles (leader or follower) in the *ACTIONS* and the *WORDS* treatment. Results of these estimations are essentially the same as the results of estimations reported in the article. Programming code, estimations' results as well as the data are available from the corresponding author upon request.

Table 2 Variable description

Explanatory variable	Description	Treatment (s)
CONST	Constant	*BASELINE*, *ACTIONS*, *WORDS*
PERIOD	Linear order effect: period from 1 to 20	*BASELINE, ACTIONS, WORDS*
OTHERS	Sum of contributions made by the other group members in the previous period (0—0; 1—€ 10; 2—€ 20)	*BASELINE, ACTIONS, WORDS*
PAYOFF	Payoff dummy: payoff in the previous period (0—if the payoff was lower than initial endowment k and 1 otherwise)	*BASELINE, ACTIONS, WORDS*
GENDER	Gender dummy: 0—male; 1—female	*BASELINE, ACTIONS, WORDS*
AGE	Age: self-reported age	*BASELINE, ACTIONS, WORDS*
INRA	Incentivized risk attitude: a scale from 0 (risk seeking) to 10 (extremely risk averse), based on the number of "safe" choices made in the Holt and Laury (2002) risk attitude elicitation procedure	*BASELINE, ACTIONS, WORDS*
SRRA	Self-reported risk attitude: self-reported individual risk attitude on a scale from 1 (very risk seeking) to 6 (very risk averse)	*BASELINE, ACTIONS, WORDS*
LANGUAGE	Language dummy: 0 – not a native speaker of the German language; 1—otherwise	*BASELINE, ACTIONS, WORDS*
MAJOR	Major dummy: 0—not a student of Economics or Business Administration; 1—otherwise	*BASELINE, ACTIONS, WORDS*
EXPERIENCE	Experience: self-reported number of times, when a participant has taken part in economic experiments before (0—never before; 1—one time; 2—from 2 to 5 times; 3—more than 5 times)	*BASELINE, ACTIONS, WORDS*
INCOME	Self-reported annual income: 1—less than € 15,000; 2—from € 15,001 to € 30,000; 3—from € 30,001 to € 45,000; 4—from € 45,001 to € 60,000; 5—more than € 60,000	*BASELINE, ACTIONS, WORDS*
SESSION	Session dummy: 0—session 1; 1—session 2	*BASELINE, ACTIONS, WORDS*
ROLE	Role dummy: 0—follower; 1—leader	*ACTIONS, WORDS*
LCONTR	Leader's contribution: amount of the leader's contribution in the current period	*ACTIONS*
LPREVCONTR	Leader's previous contribution: amount of the leader's contribution in the previous period	*ACTIONS, WORDS*

Table 2 continued

Explanatory variable	Description	Treatment(s)
NBA	Content of the non-binding announcement: 0 if $m = 0$; 1 if $m = 1$	*WORDS*
INC	Inconsistent announcement dummy: 0—in previous period the leader's announcement was consistent with her actual contribution; 1—otherwise.	*WORDS*

The content of the non-binding announcement (variable **NBA**) is one of the important factors that influence participants' decisions. Particularly, participants are more likely to contribute $c_i^t = k$ if a leader has promised to contribute k to the group activity. At the same time, the value of the leader's final contribution in the previous period (variable **LPREVCONTR**) is not a statistically significant determinant of individual behavior. This finding suggests that a non-binding announcement has a higher impact on the individual decisions than the *ex post* observation of the leader's contribution.

It may seem that participants do not condition their contributions in the current period on the previously observed leader's contribution because leaders are determined at random in every period. Since leaders change very often, followers may hope that the leader in the current period is more consistent between announced and implemented contributions than the leader in the previous period. Therefore, they may ignore the outcome of the previous period when making decisions. If this conjecture is correct, the implication is that participants should neglect the institutional reputation of leaders by not taking into account whether the leader has made a consistent or an inconsistent announcement in the previous period. However, the data fails to confirm this implication. Despite rotating leadership, participants are less likely to make positive contributions to the group activity if the leader's announcement in the previous period did not coincide with her actual decision (variable **INC**).

Note, however, that in the voluntary contribution game, not contributing is an equilibrium strategy for all players. Therefore, it is also likely that followers do not expect the leader to make a positive contribution to the group activity simply because it is irrational. Yet, it is important to them whether the leader is trustworthy or not, i.e., whether the leader's announced contribution coincides with her implemented contribution. In this case, followers are more likely to be disappointed if the leader makes an inconsistent announcement and tries to take advantage of the other group members than if the leader contributes nothing to the group activity. In order to check whether followers take into account the reputation of each individual leader in the *WORDS* treatment, we conduct an additional regression analysis.

We construct the dynamic reputation variable for each leader in the *WORDS* treatment. In the period when they are first assigned a leadership role, leaders do not have any reputation. When assigned the leadership role for the second time, the reputation variable is equal to 0 if a leader's contribution in the first period has been consistent with her announcement and 1 if it was inconsistent. In the subsequent periods

Table 3 Results of the random intercept logit regressions

Explanatory variable	Model 1			Model 2			Model 3			Model 4		
	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS
LPREVCONTR	—	-0.651** (0.227)	0.273 (0.331)	—	-0.495• (0.266)	-0.021 (0.370)	—	-0.516* (0.264)	0.006 (0.370)	—	-0.530* (0.265)	-0.010 (0.370)
NBA	—	—	1.056*** (0.300)	—	—	0.971*** (0.303)	—	—	0.996*** (0.303)	—	—	0.983*** (0.303)
INC	—	—	-0.820* (0.334)	—	—	-0.887** (0.338)	—	—	-0.843* (0.339)	—	—	-0.830* (0.342)
LCONTR	—	1.844*** (0.233)	—	—	1.819*** (0.234)	—	—	1.781*** (0.234)	—	—	1.774*** (0.234)	—
ROLE	—	-0.489* (0.239)	-1.470*** (0.240)	—	-0.520* (0.241)	-1.549*** (0.247)	—	-0.518* (0.241)	-1.545*** (0.247)	—	-0.497* (0.240)	-1.552*** (0.247)
CONST	0.932** (0.295)	0.352 (0.333)	1.330** (0.440)	-0.064 (1.443)	0.203 (0.754)	0.130 (0.944)	0.185 (1.647)	0.215 (0.874)	-0.709 (1.216)	0.504 (1.850)	2.637 (1.877)	1.016 (3.805)
PERIOD	-0.180*** (0.022)	-0.040* (0.020)	-0.077*** (0.021)	-0.174*** (0.028)	-0.044** (0.020)	-0.068** (0.022)	-0.174*** (0.028)	-0.045* (0.020)	-0.068** (0.022)	-0.175*** (0.028)	-0.046* (0.020)	-0.068** (0.022)
OTHERS				0.034 (0.215)	-0.207 (0.212)	0.459* (0.238)	0.029 (0.215)	-0.224 (0.213)	0.436• (0.239)	0.020 (0.216)	-0.226 (0.212)	0.440• (0.240)
PAYOFF				0.317 (0.399)	-0.176 (0.457)	-0.797 (0.528)	0.323 (0.399)	-0.168 (0.456)	-0.785 (0.528)	0.340 (0.400)	-0.164 (0.457)	-0.794 (0.530)
GENDER				0.670 (0.444)	-0.050 (0.315)	0.504 (0.364)	0.746 (0.492)	-0.137 (0.292)	0.566 (0.365)	0.618 (0.504)	-0.075 (0.304)	0.250 (0.404)

Table 3 continued

Explanatory variable	Marginal effect (standard error)											
	Model 1			Model 2			Model 3			Model 4		
	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS	BASELINE	ACTIONS	WORDS
AGE				0.008	0.021	0.047	0.010	0.032	0.038	0.016	0.018	0.043
				(0.051)	(0.023)	(0.028)	(0.051)	(0.022)	(0.028)	(0.050)	(0.022)	(0.030)
INRA							0.034	−0.131	0.112	0.032	−0.124	0.116
							(0.153)	(0.089)	(0.125)	(0.161)	(0.097)	(0.159)
SRRA							−0.138	0.164	0.105	−0.197	0.082	0.037
							(0.229)	(0.123)	(0.169)	(0.244)	(0.131)	(0.200)
LANGUAGE										−0.094	0.118	0.206
										(0.725)	(0.438)	(0.687)
MAJOR										0.806	−0.135	−0.720
										(0.882)	(0.380)	(0.483)
EXPERIENCE										−0.132	−0.254	−0.051
										(0.245)	(0.199)	(0.203)
INCOME										−0.307	−0.317	−1.007
										(0.507)	(0.466)	(0.838)
SESSION										0.235	−0.367	−0.065
										(0.492)	(0.335)	(0.530)
LL	−254.816	−265.111	−251.457	−240.752	−263.737	−247.251	−240.567	−261.554	−246.656	−239.630	−259.442	−244.989
Individual effects[a]	0.931	0.604	0.768	0.848	0.553	0.674	0.845	0.451	0.648	0.801	0.360	0.569
	(0.184)	(0.060)	(0.108)	(0.152)	(0.050)	(0.245)	(0.150)	(0.034)	(0.076)	(0.133)	(0.023)	(0.057)
N (decisions)	456	456	456	456	456	456	456	456	456	456	456	456
N (individuals)	24	24	24	24	24	24	24	24	24	24	24	24

• Marginally significant at $p < 0.07$

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

[a] Standard deviation (standard error)

when assigned the leadership role, 1 (0) is added to the leader's reputation variable if she has made an inconsistent (a consistent) announcement in the previous period. We use ordinary least squares regression (with controls for the time period of the game) to explore whether followers take into account leader's reputation. We find that the better is a leader's reputation (the fewer inconsistent announcements this leader makes as the game progresses) the higher is the sum of follower's contributions ($p < 0.004$, $R^2 = 0.171$). Therefore, followers indeed pay attention to the consistency of leaders' announcements.

Table 3 also reports two interesting cross-treatment effects. First, participants in the *ACTIONS* treatment do not take into account whether or not other group members have made positive contributions in the previous period (variable **OTHERS**). However, this variable is significant in the *WORDS* treatment: players are more likely to make positive contributions in the current period if they have observed that other players have contributed positive amounts in the previous period. Second, while leaders' contributions in the previous period (variable **LPREVCONTR**) do not affect current contributions in the *WORDS* treatment, this variable is significant in the *ACTIONS* treatment. Participants in the *ACTIONS* treatment are more likely to contribute $c_i^t = k$ to the group activity in the current period after observing that the leader has contributed nothing in the previous period. These cross-treatment effects suggest that a non-binding announcement and a binding commitment have a different impact on individual behavior in the voluntary contribution game.

In order to explore the impact of **OTHERS** and **LPREVCONTR** in more detail, we conduct additional random intercept logit regressions. We check whether these two variables have different effects on participants' decisions when they play roles of leaders and followers. Results of these additional estimations are reported in Table 4.

According to Table 4, when participants play the roles of followers in the *ACTIONS* treatment, they are more likely to contribute to the group activity if they have observed a zero leader's contribution and a low sum of other group members' contributions in the previous period. In contrast, **OTHERS** and **LPREVCONTR** are not significant determinants of participants' decisions if they play the roles of leaders in the *ACTIONS* treatment. However, in the *WORDS* treatment, participants are more likely to contribute positive amounts to the group activity after observing a relatively high sum of others' contributions in the previous period, irrespective of the role. Yet, while **LPREVCONTR** is not an important determinant of the followers' behavior in the *WORDS* treatment, after observing a relatively high leader's final contribution in the previous period, the leader in the current period is more likely to contribute a positive amount to the group activity. Therefore, the data suggest that participants are more likely to focus on reaching social efficiency in *ACTIONS* than in *WORDS*. On the one hand, followers in the *ACTIONS* treatment engage in facilitating the attainment of the social optimum. They do so by making positive contributions to the group activity when the leader fails to set a good example as well as when they observe a relatively low sum of others' contributions in the previous period. On the other hand, followers in the *WORDS* treatment are less likely to contribute after observing a relatively low sum of others' contributions in the previous period.

In all estimations, we control for individual unobserved heterogeneity of the experimental participants. By incorporating unobserved heterogeneity into our analysis, we

Table 4 The impact of other group members' contributions and leader's final contribution on individual decisions by role

Explanatory variable	Treatment			
	ACTIONS		*WORDS*	
	Leaders	Followers	Leaders	Followers
Random intercept logit regression 1				
CONST	0.284	0.908***	−1.795*	0.219
	(0.421)	(0.260)	(0.709)	(0.283)
OTHERS	−0.166	−0.411*	1.028**	0.625**
	(0.267)	(0.173)	(0.364)	(0.205)
Log-likelihood (LL)	−99.814	−200.299	−85.480	−175.499
Standard deviation (standard error) for the random intercept (level 2)	1.031 (0.349)	0.455 (0.040)	2.048 (2.258)	0.442 (0.043)
Number of level 1 units (contribution decision)	152	304	152	304
Number of level 2 units (individual)	24	24	24	24
Random intercept logit regression 2				
CONST	0.018	0.800***	−1.084*	0.742***
	(0.354)	(0.214)	(0.459)	(0.183)
LPREVCONTR	0.138	−0.696**	1.585***	0.539
	(0.384)	(0.251)	(0.462)	(0.277)
Log-likelihood (LL)	−99.943	−199.243	−84.034	−178.454
Standard deviation (standard error) for the random intercept (level 2)	1.022 (0.344)	0.491 (0.046)	1.637 (1.219)	0.338 (0.028)
Number of level 1 units (contribution decision)	152	304	152	304
Number of level 2 units (individual)	24	24	24	24

insure that regression results are reliable and robust and obtained estimates are unbiased and efficient. Tables 3 and 4 suggest that in all treatments unobserved individual heterogeneity has an important impact on contributions. Particularly, the standard deviation of the random intercept at the level of individual participants is greater than 0.35 in all estimations. This finding indicates that apart from factors measured in the experiment, other individual characteristics such as, e.g., cultural socio-economic and psychological parameters may have an impact on individual contributions. Developing efficient techniques which would allow measuring with high degree of precision a large menu of possible determinants of individual decisions in the laboratory (through incentivized procedures as well as questionnaires) is an important endeavor for the future research in economics and psychology.

Even though our voluntary contribution game has only one Nash equilibrium, experimental participants face strategic risk which may influence their decisions. Particularly, a participant who is averse to strategic risk may contribute nothing to the group activity simply due to her strategic risk aversion. In its turn, strategic risk attitude may be correlated with individual risk attitude. In this case, individual risk attitude may be used a proxy of strategic risk attitude. In order to make sure that risk taking preferences

Table 5 Risk attitudes of experimental participants

Constant relative risk aversion (CRRA) characteristic			Number of participants (%)		
Risk attitude rank[a]	CRRA coefficient r	Description	*BASELINE*	*ACTIONS*	*WORDS*
0–1	$r < -0.95$	Highly risk seeking	1 (4.2)	0 (0.0)	0 (0.0)
2	$-0.95 < r \le -0.49$	Very risk seeking	0 (0.0)	0(0.0)	1 (4.2)
3	$-0.49 < r \le -0.15$	Risk seeking	0 (0.0)	2 (8.3)	1 (4.2)
4	$-0.15 < r \le 0.15$	Risk neutral	1 (4.2)	5 (20.8)	5 (20.8)
5	$0.15 < r \le 0.41$	Slightly risk averse	4 (16.7)	4(16.7)	4 (16.7)
6	$0.41 < r \le 0.68$	Risk averse	7 (29.2)	2 (8.3)	7 (29.2)
7	$0.68 < r \le 0.97$	Very risk averse	3 (12.5)	4(16.7)	1 (4.2)
8	$0.97 < r \le 1.37$	Highly risk averse	3 (12.5)	0(0.0)	1 (4.2)
9 or 10	$r > 1.37$	Stay in bed	0 (0.0)	2 (8.3)	1 (4.2)
Kruskal–Wallis equality-of-populations rank test				$p = 0.2856$	
Average rank			5.8	5.6	5.3
Median rank			6	5	5
Standard deviation			1.8	2.0	1.6
Inconsistent			5 (20.8)	5 (20.8)	3 (12.5)

In the econometric analysis, inconsistent subjects were assigned a median rank (6 in the *BASELINE* treatment and 5 in the *ACTIONS* and the *WORDS* treatment)

[a] Number of "safe" choices made in the Holt and Laury (2002) procedure

do not intervene with the participants' decision making we control for risk attitudes in our econometric analysis.

Recall from Sect. 2 that in the second experimental task, participants have taken part in the Holt and Laury (2002) risk attitude elicitation procedure. This procedure offers ten consecutive pairwise choices between a relatively safe and a relatively risky lottery. The probabilities of payoffs in each of these two lotteries are varied in such a way that at some point, an individual should switch from opting for a relatively safe lottery to a relatively risky lottery. The number of "safe" choices made before this switch point is often used as a proxy of an individual's risk attitude.

According to the procedure, more than half (59.7%) of the participants in our experiment are at least slightly risk averse. The average risk attitude rank in the experiment is 5.6 with the median of 6 and a standard deviation of 1.8. Table 5 provides a cross-treatment comparison of risk attitude ranks. According to the results of the Kruskal–Wallis equality of populations rank test with two degrees of freedom, experimental participants are homogeneous in terms of their risk attitudes in all three treatments ($p > 0.20$).

In addition to the Holt and Laury (2002) procedure, we ask participants to indicate their attitude toward risk in the post-experimental questionnaire. We include an indicator of an individual risk attitude obtained from the second experimental task (variable **INRA**) as well as the self-reported measure of the risk attitude (variable **SRRA**) in all estimations. Notably, neither of these two measures appear to be statistically significant in any of the estimated models.

4 Discussion and conclusion

This article has compared words and actions as a means to influence cooperation in a social dilemma. We consider a simple voluntary contribution game where a leader can influence the group cooperation either by making a non-binding announcement of her intended contribution or by contributing first to the group activity. Our results suggest that words and actions increase the level of individual contributions to a similar extent. Furthermore, a leader's non-binding announcement and binding commitment generate a statistically significantly higher level of cooperation than is reached in the control treatment where all players make simultaneous decisions. This finding suggests that high levels of cooperation can be reached by means of communication rather than taking action.

We also observe that followers react differently to leaders' words and actions. Particularly, they tend to contribute higher amounts after observing a non-binding announcement than a binding commitment. This result can be explained by the fact that announcements are systematically larger than commitments.

Our results suggest that participants are more likely to focus on reaching social efficiency after observing a leader's binding contributions. It appears that when intentions of the leader are observable and the leader fails to set a good example, in the next period, followers take on the leadership role and try to achieve higher payoffs without relying on the leader. However, when the intentions of the leader are unobservable, a relatively low sum of other players' contributions in the previous period (especially when the leader makes an announcement inconsistent with her subsequent decision) has a negative impact on followers' desire to attain a social optimum. Instead, they focus on preserving their endowments.

We also find that imposed roles (forced identity) have a significant impact on behavior. Research in social psychology and sociology has indicated that not only do individuals with similar characteristics behave differently when they are assigned different roles, but also the same individual may exhibit different patterns of behaviors when assigned different roles (e.g., Pollay 1968; Callero 1994). This article relates this psychological literature on forced identity with the economics research on voluntary contribution games.

In our experiment, individuals exhibit different patterns of behaviors in a situation when they play the roles of leaders compared with a situation where they play the roles of followers. This result has important implications for research on group cooperation in social dilemmas. One possibility is that observed behavior can be rationalized by the quantal response logic (e.g., Offerman et al. 1998) which has a different error structure for leaders and followers. Uncovering the relative impact of forced identity on the propensity to contribute in a voluntary contribution game is an important endeavor for future research.

Recent developments in decision theory emphasize the importance of group goals and social motives on individual decision making (e.g., Krantz and Kunreuther 2007). In laboratory experiments, Charness et al. (2007) and Chen and Li (2008) find that group membership influences individual behavior in many ways. Our results contribute to this stream of literature by suggesting that not only group membership per se, but also the context in which this group operates and a decision task that this group

faces, has an impact on individual decision making. When the leader makes a commitment instead of simply announcing her intentions to the rest of the group, followers become more engaged in group activity as well as more interested in reaching the social optimum.

In the treatment with a non-binding announcement, failure to fulfill the announced plan not only tends to hurt the reputation of a current leader, but also negatively influences institutional reputation of all subsequent leaders. Particularly, when participants observe that a leader has not contributed the announced amount in the previous period, they decrease the level of their contributions even if the leader has been re-assigned at the beginning of the current period. This finding indicates that unfulfilled announcements create an erosion of trust at the institutional level. Such announcements discourage followers from cooperating even if the leaders are reassigned every period. Exploring the robustness of this finding is an important direction for the future research agenda of the literature on voluntary contribution games. Other possible extensions of our research include relaxing the restriction on individual contributions, varying the context and the content of non-binding announcements and allowing leaders to endogenously decide between using words and actions to influence cooperation.

References

Arbak, E., & Villeval, M. (2007). Endogenous leadership selection and influence. Université Lyon 2 Working Paper.

Andreoni, J., & Croson, R. (2008). Partners versus strangers: Random rematching in public goods experiments. In C. R. Plott & V. L. Smith (Eds.), *Handbook of experimental economics results*, Chap. 82 (pp. 776–783). Elsevier, Amsterdam.

Baliga, S., & Morris, S. (2002). Co-ordination, spillovers, and pre-game communication. *Journal of Economic Theory, 105*(2), 450–468.

Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution, 54*, 39–57.

Blume, A., & Ortmann, A. (2007). The effects of costless pre-play communication: Experimental evidence from games with pareto-ranked equilibria. *Journal of Economic Theory, 132*(1), 274–290.

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

Callero, P. (1994). From role playing to role-using: Understanding role as resource. *Social Psychology Quarterly, 57*(3), 228–243.

Charness, G., & Grosskopf, B. (2004). What makes pre-game communication effective? Experimental evidence. *Economics Letters, 83*(3), 383–389.

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review, 97*(4), 1340–1352.

Chen, Y., & Li, S. X. (2008). Group identity and social preferences. *American Economic Review, 99*(1), 431–457.

Crawford, V. (1998). A survey of experiments on communication via pre-game communication. *Journal of Economic Theory, 78*(2), 286–298.

Crawford, V., & Sobel, J. (1982). Strategic information transmission. *Econometrica, 50*(6), 1431–1451.

Farrell, J. (1987). Pre-game communication, coordination, and entry. *RAND Journal of Economics, 18*(1), 34–39.

Farrell, J. (1988). Communication, coordination and Nash equilibrium. *Economic Letters, 27*(3), 209–215.

Farrell, J., & Rabin, M. (1996). Pre-game communication. *Journal of Economic Perspectives, 10*(3), 103–118.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178.

Gächter, S., Nosenzo, D., Renner, E., & Sefton, M. (2010). Who makes a good leader? Cooperativeness, optimism and leading-by-example *Economic Inquiry* (forthcoming).

Gächter, S., & Renner, E. (2003). Leading by example in the presence of free rider incentives, University of St. Gallen Working Paper.

Genesove, D., & Mullin, W. P. (2001). Rules, communication and collusion: Narrative evidence from the sugar institute case. *American Economic Review, 91*(3), 379–398.

Green, J. R., & Stokey, N. L. (1980). A two-person game of information transmission, Discussion Papers 418, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Güth, W., Levati, V., Sutter, M., & Vander Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics, 91*(5–6), 1023–1042.

Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *American Economic Review, 88*(5), 1188–1206.

Holt, Ch., & Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review, 92*(5), 1644–1655.

Isaac, R. M., Walker, J., & Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice, 43*(2), 113–149.

Krantz, D. H., & Kunreuther, H. C. (2007). Goals and plans in decision making. *Judgment and Decision Making, 2*(3), 137–168.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621.

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

Longford, N. T. (1994). Logistic regression with random coefficients. *Computational Statistics and Data Analysis, 17*(1), 1–15.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*, 50–60.

Moxnes, E., & Vander Heijden, E. (2003). The effect of leadership in a public bad experiment. *Journal of Conflict Resolution, 47*(6), 773–795.

Muller, L., Sefton, M., Steinberg, R., & Vesterlund, L. (2008). Strategic behavior and learning in repeated voluntary contribution experiments. *Journal of Economic Behavior and Organization, 67*(3–4), 782–793.

Offerman, Th., Schram, A., & Sonnemans, J. (1998). Quantal response models in step-level public good games. *European Journal of Political Economy, 14*(1), 89–100.

Ottaviani, M., & Sørensen, P. N. (2006). Reputational pre-game communication. *RAND Journal of Economics, 37*(1), 155–175.

Palfrey, T. R., & Prisbrey, J. E. (1997). Anomalous behavior in public goods experiments: How much and why?. *American Economic Review, 87*, 829–846.

Pollay, R. (1968). The measurement of role differentiation. *Sociometry, 31*(1), 120–124.

Potters, J., Sefton, M., & Vesterlund, L. (2007). Leading-by-example and signaling in voluntary contribution games: An experimental study. *Economic Theory, 33*(1), 169–182.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2002). Reliable estimation of generalised linear mixed models using adaptive quadrature. *Stata Journal, 2*(1), 1–21.

Rabin, M. (1994). A model of pre-game communication. *Journal of Economic Theory, 63*(2), 370–391.

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

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83.