CEPA DP No. 73

FEBRUAR 2024

Pre-election communication in public good games with endogenous leaders

Lisa Bruttel
Gerald Eisenkopf
Juri Nithammer

Center for Economic Policy Analysis

CEPA Discussion Papers

Center for Economic Policy Analysis
https://www.uni-potsdam.de/cepa

University of Potsdam
August-Bebel-Straße 89, 14482 Potsdam
Tel.: +49 331 977-3225
Fax: +49 331 977-3210
E-Mail: dp-cepa@uni-potsdam.de

ISSN (online) 2628-653X

CEPA Discussion Papers can be downloaded from RePEc
https://ideas.repec.org/s/pot/cepadp.html

Published online at the Institutional Repository of the University of Potsdam
https://doi.org/10.25932/publishup-62395

CEPA Discussion Paper No. 73 Februar 2024

Pre-Election Communication in Public Good Games with Endogenous Leaders*

Lisa Bruttel
University of Potsdam

Gerald Eisenkopf
University of Vechta

Juri Nithammer
University of Potsdam

ABSTRACT

Leadership plays an important role for the efficient and fair solution of social dilemmas but the effectiveness of a leader can vary substantially. Two main factors of leadership impact are the ability to induce high contributions by all group members and the (expected) fair use of power. Participants in our experiment decide about contributions to a public good. After all contributions are made, the leader can choose how much of the joint earnings to assign to herself; the remainder is distributed equally among the followers. Using machine learning techniques, we study whether the content of initial open statements by the group members predicts their behavior as a leader and whether groups are able to identify such clues and endogenously appoint a "good" leader to solve the dilemma. We find that leaders who promise fairness are more likely to behave fairly, and that followers appoint as leaders those who write more explicitly about fairness and efficiency. However, in their contribution decision, followers focus on the leader's first-move contribution and place less importance on the content of the leader's statements.

Keywords: Leadership, Public good, Voting, Promises, Experiment
JEL Codes: C92, D23, D72, D83

Corresponding author:
Juri Nithammer
University of Potsdam
August-Bebel-Straße 89
14482 Potsdam, Germany
Email: juri.nithammer@uni-potsdam.de

*Experimental design and procedures, number of subjects, hypotheses and main data analysis specifications were preregistered at OSF before the data collection had started (osf.io/jh42s/). Declaration of interests: None. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1 Introduction

The appointment of effective and trustworthy leaders plays an important role for enhancing cooperation within groups. A vast literature shows that having a leader can increase total contributions of group members to a public good (see, e.g. Gächter and Renner, 2018, and the literature therein). However, the effectiveness of a leader can vary substantially, in particular when leadership includes control over the distribution of common resources. Such a powerful leader has strong incentives to maximize contributions to the public good. However, such power also allows for abuse because the leader can disproportionately allocate these resources for the own private benefit. Hence, the position of the leader is attractive for both pro-social as well as purely selfish group members. The so-called "resource case" illustrates the dilemma, as some leaders in resource-rich economies undermine political institutions to extract more rents from the resources with detrimental effects on efficiency and equality (Caselli and Cunningham, 2009; Ross, 2015).

Thus, groups have an interest to screen leaders in advance and elect someone they think will act in a way that fosters both efficiency (by motivating the group members to contribute to the public good) and fairness (by using the leadership power in a non-selfish way). However, as the groups' information about the candidates' types is likely to be limited, they may fail to appoint such a "good" leader and end up with a "bad" one who does not satisfy one or both criteria. In this paper, we analyze how communication of leadership candidates improves appointment choices, increases contributions of fellow group members and indicates fairness in subsequent distributions.

Our experimental design replicates the main features of the above-described situation: the group members invest private resources into a public good. Any contribution is efficiency-enhancing. Ahead of the contributions, they elect one of them as the leader. The leader makes a public contribution decision before the other group members decide. Furthermore, the leader has distribution power, i.e., she can transfer resources from the public good into her own account. The remainder is distributed equally among the members of the group.

This allocation decision allows for an abuse of power; no contract forces the leader to implement a specific distribution. Hence, the group members have an incentive to select their leader carefully. However, they only get an unreliable signal about the trustworthiness of the leader: all group members formulate written statements on how they plan to act as a leader, which are displayed to everyone in the group before the election.

For the analysis of the statements, we use machine learning techniques to quantify the general quality of the statements and how specific they are about the two concepts efficiency and fairness. We find that high quality statements receive more votes in the election. However, they have almost no impact on the contributions of the followers, which are affected by the leaders' first-move contribution decisions. Regarding leadership behavior, we find that the more a statement is specific about fairness, the higher is the share of the joint benefit the leader assigns to the followers.

In the following, we summarize the relevant literature in Section 2, describe our experimental design in Section 3, and develop hypotheses in Section 4. We then present the experimental results in Section 5 and conclude in Section 6. An appendix complements the paper with additional results and a translation of the instructions.

2 Contribution to the Literature

Our study focuses on the messages ahead of the appointment of the leader. The subsequent game with leadership represents an asymmetric social dilemma. In a symmetric social dilemma like the standard public good games (e.g. Ledyard, 1995; Fischbacher et al., 2001; Fischbacher and Gächter, 2010), all members of a group make simultaneous decisions and face the same trade-off between providing for a pareto superior outcome and maximizing their own payoff. A vast amount of literature shows that many people contribute more than they would in the non-cooperative Nash equilibrium, in particular if others do so as well (see, for example, the meta-analyses by Zelmer, 2003; Fiala and Suetens, 2017; Spadaro et al., 2022).

Leading-by-example introduces an asymmetry into this social dilemma because one group member decides first. The other group members observe this decision before they make their own choice (Güth et al., 2007). This sequential decision making tends to increase aggregate contributions but often implies a 'leader's curse' because the second movers (or followers) make, on average, lower contributions (Gächter and Renner, 2018; Eichenseer, 2019; Eisenkopf and Kölpin, 2021). Hence, while the appointment of a leader can generate a positive external effect for the fellow group members, most experimental participants consider the position itself as rather undesirable (Haigner and Wakolbinger, 2010; Arbak and Villeval, 2013; Cappelen et al., 2016). It is therefore unsurprising that the pro-socially minded have a higher inclination to lead by example (Bruttel and Fischbacher, 2013; Centorrino and Concina, 2013; Préget et al., 2016). Those who volunteer to become leaders tend to contribute more than those who have been forced into that role (Haigner and Wakolbinger, 2010; Rivas and Sutter, 2011; Dannenberg, 2015).

The asymmetry in the social dilemma changes if one allocates additional property rights to the leadership. In our case, the leader can control the distribution of the joint benefit. This allocation right for the leader represents a much stronger impact than the lump sum compensation for leaders introduced by Cappelen et al. (2016). While such a compensation induces less pro-socially inclined group members to become leaders it does not change the incentive structure during the subsequent game. Interestingly, Cappelen et al. (2016) report that high compensations for leaders are detrimental for public good provision which indicates that outcome and kindness related considerations for reciprocity are less relevant (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006) in this context.

Several studies have investigated the impact of individualized allocation power in simultaneous public good games (i.e., without leading by example). Van der Heijden et al. (2009) find that most people with such power forego the temptation to appropriate team output. As a result, compared to revenue-sharing, the presence of a team leader results in a significant improvement in team performance. Nevertheless, Drouvelis et al. (2017) observe in a comparable setting that teams tend to resist the allocation of distributive power to a player. Similarly, Stoddard et al. (2021) observe a trade-off between efficiency and moral hazard. They show that allocators with the most enhanced flexibility in distributions generate the highest contributions among the group members but produce a particularly high variance in allocation decisions. Karakostas et al. (2023) compare such allocation rights across different social dilemmas and observe a positive impact on contributions in linear and best-shot public-good games but not in weakest-link games.

Apart from focusing on simultaneous games, it is important to highlight another difference between our study and these papers on individualized allocation power. In our case, the group members make independent contribution decisions but receive a one-size-fits-all transfer. This is a novel feature in the literature because it investigates the reciprocity of people towards entire groups. Eisenkopf and Walter (2021) study the aforementioned leadership game with punishment options in which some leaders can target individuals with their punishment while others cannot differentiate between the followers. The latter setting induces less cooperation as it imposes a greater risk of arbitrary punishment. Song (2008) distinguishes between trustors who decide either for themselves or for an entire group. The results show that acting as a representative reduces trust.

Overall, the evidence in the literature suggests that contributions depend on the trust of the group members in their leader. Leaders can increase trust with their initial messages. After all, communication is an ubiquitous tool to coordinate and motivate groups. Brandts et al. (2015) report experimental results on leadership effects in a turnaround game. They show that communication from leaders had a greater effect than incentives for followers. Some studies even find that simple contribution recommendations in a public good game can substitute leading-by example to a large extent (Pogrebna et al., 2011; Koukoumelis et al., 2012; Eisenkopf, 2020; Eisenkopf and Walter, 2021) but other papers cannot confirm this effect (Dannenberg, 2015; Eichenseer, 2019). Feltovich and Grossman (2015) observe that the impact decreases with group size in a threshold public-good game.

A key problem of the leader is how to signal her trustworthiness. Ben-Ner et al. (2011) show that communication tends to increase trust and trustworthiness in a trust game and that the content of the messages matters. Nielsen et al. (2019) replicate this result for a trust game between individuals but fail to do so if the trustee consists of a team. The increased cooperation rests on the expectation that the other party sticks to its promises. The literature on promise-keeping and guilt-aversion has shown that such statements are more than just cheap talk (Vanberg, 2008; Rosenbaum et al., 2014; Bellemare et al., 2019), so that we can assume that they are to some extent informative about the leader's type. However, selfish leaders with a low level of guilt aversion might find it optimal to lie and pretend they would use the leader's power in a fair way while planning to grab the whole joint benefit if being elected the leader. Eisenkopf and Bächtiger (2013) show with a common pool game with ex-ante communication that false promises can lead to particularly large inequality in outcomes, as the deceived parties contribute substantially to the joint benefit without getting much out of it.

3 Experimental Design

Participants in our experiment interact in fixed, randomly allocated groups of four players, where one is the leader and the other three are followers. We name the roles neutrally, with the leader being in "position X" and the followers in "position Y."

The Game The main game is a one-shot social dilemma which is framed in terms of contributions to a public good. All players receive an endowment of 100 points and can decide how much of this endowment to contribute. The total sum of contributions is multiplied by the factor 1.6. The game is played using the strategy method. Each participant makes all decisions listed below and learns her role only at the very end of the sequence of decisions.

In stage 1, participants formulate a statement how they will act in position X and how they recommend the others to act in position Y. Afterwards, all four statements are displayed to all participants in the same group. Based on these statements, the group members elect the leader in stage 2. In order to elicit their true preferences, we use a variant of the random dictator procedure. The group members assign six votes to the four group members, with no ties allowed. Each vote is worth 1/6 probability to be elected as the leader. After all group members stated their preferences, one of them is selected randomly as the "voting dictator" and his or her lottery is played out to determine the leader. Importantly, at this stage, participants are not informed on who was elected to be the leader so that we can elicit contribution decisions using the strategy method.

In the next two stages, all group members state their contribution decisions for both roles. In stage 3, they state their unconditional contribution which would be implemented if they were elected the leader. Note that in our setup it is strictly optimal to contribute the entire endowment as the leader has unlimited distribution power over the public good, and thus, faces no strategic risk. In stage 4, participants decide on their conditional contributions given the leader statements and leader contributions of the other three group members.

In stage 5, each player decides how much of the public good to assign to herself given that she is elected the leader. The rest is divided equally among the other three group members. Only in the final, 6th stage, the outcome of the role assignment procedure in stage 2 is revealed to the participants. They receive feedback on the resulting contributions and the final payoffs of all group members.

All participants in this experiment face the same treatment. We use endogenous differences in the initial statements to test our hypotheses.

Procedures We collected the data in the laboratory at the University of Potsdam PLEx. The participants were invited for the sessions through the regular invitation procedures of the laboratory. No specific rules have been used to restrict subjects that are registered in the database from participating in the experiment other than that they have not participated in this experimental study before. We programmed the experiment in z-Tree (Fischbacher, 2007) and invited subjects using ORSEE (Greiner, 2015). All participants signed informed-consent sheets before the sessions started.

Our sample size is 144 participants, which we collected in sessions with a minimum of 8 subjects per session. They received a show-up-fee of 10 euros. We employ a conversion rate of 16/1 of points to euros. If all group members contribute their entire endowment of 100 points, and the leader distributes the joint benefit evenly, each group member receives a profit of 160 points, i.e. 10 euros.

4 Hypotheses

Our experimental design represents a combination of a public good game and a trust game, enhanced with the endogenous leadership assignment. As in a public good game, all group members can make contributions. For conditional cooperators, the contribution depends on their beliefs about the contributions of the other group members. As in the trust game, the group members have to trust the leader that she will not appropriate the returns from contributions disproportionally for her own benefit.

Standard game theory suggests, via backward induction, that followers should not make any contribution because the leader maximizes her own payoff by assigning the whole joint benefit to herself. The leader, in turn, should contribute everything, benefit from the efficiency of the public good, and then keep everything. However, communication typically enhances both trust and trustworthiness (Fiedler and Haruvy, 2009; Ben-Ner et al., 2011; Sheremeta and Zhang, 2014), which may induce positive contributions of the followers and not perfectly selfish decisions of the leader. Furthermore, the leader's contributions may still provide a role model that encourages contributions from the followers (Güth et al., 2007; Eisenkopf, 2020).

Our analysis focuses on how differences in the initial statements relate to subsequent election decisions, contributions of followers and allocation decisions of the leader. We expect that the statements will vary with respect to how they emphasize the two criteria efficiency and fairness. Efficiency relates to the argument that the joint benefit gets larger if the group members invest more. It is therefore likely to be a main argument regarding the behavior expected by the followers. Fairness is the key element of any leader's promise to reward the group members for their contributions instead of assigning the entire public good to herself.

Our endogenous leadership selection process with a differentiated assignment of lottery tickets by a random dictator eliminates strategic voting. Hence, each group member can rank all others as leaders according to how she assesses their initial statements. While the rank assigned to herself indicates that group member's relative preference for leadership, the ranks of the fellow group members indicate their respective trustworthiness (i.e. whether they will reward a high contribution of the followers).[1]

Given that the group members have an interest in electing a good leader, we expect that their voting decision will respond to any content in the statements that indicates that the leader will successfully coordinate the group on a high contribution level and distribute the revenue in a fair way. Furthermore, we expect that reminding the group members of the fact that the total pie gets larger if the endowments are invested in the public good as well as promises of implementing a fair distribution will increase contributions.

Hypothesis 1. *Group members who address (a) efficiency and (b) fairness receive more votes.*

Hypothesis 2. *Group members who address (a) efficiency and (b) fairness induce larger contributions by the followers.*

From the literature on promise-keeping and guilt-aversion we derive that there is a cost of lying. Thus, we expect that leader statements are predictive for their behavior to some extent. Leaders promising to distribute the public good in a fair way will be more likely to do so than leaders who did not make such a promise.

Hypothesis 3. *Group members who address fairness also make fairer distributions.*

Regarding the operationalization of fairness in distributions we rely on two different definitions. Let c_L denote the contribution of the leader and $c_{F_{max}}$ ($c_{F_{med}}$, $c_{F_{min}}$) the highest (median, lowest) contribution of a follower, with $c_{F_{max}} \geq c_{F_{med}} \geq c_{F_{min}}$. Regarding the

[1]In our analysis, we will derive "relative ranks" from the votes. Each group member considers the messages of the other three group members and ranks them according to her preferences. She then reveals these preferences by assigning votes. The more votes she assigns, the higher her preference for this individual. Based on this, the individual to whom she assigns the most votes receives a relative rank of 3. The second of 2, and the third of 1. Thus, a higher relative rank is better. This ranking system is "relative" in the sense that the number of votes an individual assigns to herself is taken out of the ranking.

subsequent distribution of the public good, T_L denotes the transfer by the leader into her own account while T_F denotes the transfer into the account of each follower, with $T_L + 3T_F = 1.6 \cdot (c_L + c_{F_{max}} + c_{F_{med}} + c_{F_{min}})$.

The first definition of fairness names as leader fair if her share of the public good is equal to or lower than her share of contributions, i.e. $\frac{T_L}{T_L+3T_F} \leq \frac{c_L}{c_L+c_{F_{max}}+c_{F_{med}}+c_{F_{min}}}$, and unfair otherwise. This category includes both meritocratic and egalitarian distribution choices if the leader's contribution is at least as high as the average contribution of followers in her group.

As a second definition of fairness, we consider a leader fair if she compensates (at least on average) the followers for their contributions, i.e. $T_F \geq \frac{c_{F_{max}}+c_{F_{med}}+c_{F_{min}}}{3}$, and unfair otherwise. Note that this second definition allows a "fair" leader to assign a larger share of the public good to herself than the first definition.

5 Results

Our study focuses on the impact of written messages of potential leaders on the subsequent voting behavior of group members, their contributions to the public good, and the allocation of the joint benefit. Hence, before we provide statistical information about behavioral outcomes, we have to analyze the text messages in order to generate explanatory variables.

5.1 Text Analysis

We use unsupervised machine learning methods to capture the content of a given message. Such an approach requires a representation of the data contained in the texts. We use the most straight forward representation, which is a "bag of words" model. This assumes that each document (the statements, in our case) is just a collection of tokens (words, numbers...) where the ordering of the tokens does not matter. Thus, after some initial preprocessing steps, which includes manual error-correction, the removal of special characters and stop words as well as lemmatization of the statements, we transform the statements into a document-term matrix. This is a two-dimensional matrix that counts, for each document, the number of occurrences of a specific token in that document. Based on this, we run two types of analyses: unsupervised topic-modeling and simple word-counts.

We use latent dirichlet allocation to produce a topic model of the statements. Topic models assume that documents consist of topics and that topics consist of tokens. See Andres et al. (2023) for a detailed explanation of how topic modelling, via latent dirichlet allocation, can be applied in experimental economics. For our dataset, the topic model produces two separate topics. In one topic, the tokens relate to a very explicit plan of action. The ten most frequent tokens in this explicit topic are 'point,' '100,' 'position x,' 'quarter,' 'get,' 'equal,' 'contribute,' '160,' 'be' and 'group-income.' In the other topic, the leader seems to remain rather unspecific about the planned behavior. The ten most frequent tokens in this topic are 'fair,' 'position x,' 'position y,' 'end,' 'high,' 'value,' 'distribute,' 'equal,' 'determine,' and 'group-income.' Although there is overlap, the first topic would allow to make explicit that one expects contributions of '100 points' rather than just 'high' contributions, and that one would distribute profits in a way that each player gets a 'quarter' rather than just a 'fair' amount. Thus, we characterize the statements on a single dimension and interpret the share of a statement that the algorithm attributes to the explicit topic as our explicitness measure.

As the output of the unsupervised topic model is best interpreted as some measure of the general quality or explicitness of a statement, we complement this approach with a simple word count that allows a more detailed analysis of the statements based on our hypotheses. For this complementary approach, we count separately the number of occurrences of tokens that relate to a concept of fairness, to a concept of efficiency, and to a concept of modality and conditionality.[2] We select the tokens for the word count based on a list of the 100 most used tokens in the corpus and then assigned them manually based on the context of the game. For example, the token 'quarter' relates to fairness because it

[2]We include modality and conditionality because it has been shown that people may hide bad intentions behind such vague communication (Serra-Garcia et al., 2011). In our case, some participants make statements like "I would distribute the profits equally, but only if everybody contributes 100." Thus, leaders that make conditional promises of fairness are likely to behave differently (and might provoke different behavior by the followers) than leaders who made unconditional promises of fairness.

most likely refers to a leader's intent to assign a quarter of the eventual group-income to each of the members of the group. For the fairness concept, these tokens include 'equal,' 'quarter,' 'equitable,' 'divide,' 'distribute,' 'share,' 'share equally,' 'portion,' 'distribution.' For the efficiency concept, these tokens include '100,' '160,' 'one-comma-six,' '640,' 'group-account,' 'group-income.' For the concept of modality and conditionality, these tokens include 'if/when' and 'would.' [3]

While the explicitness measure provides a combined estimate to which degree a statement is explicit about both, efficiency and fairness, the word count allows a separate analysis of how followers react to efficiency arguments and to fairness arguments. Thus, for the test of Hypotheses 1 and 2, which make a joint prediction for efficiency and fairness content, we can use both, the explicitness measure and the word count. For Hypothesis 3, which relates to fairness content only, we restrict the analysis to the word count measure.

5.2 Hypotheses Tests

Table 1 shows summary statistics for the relevant variables for our analysis. The top rows present the behavioral variables, the lower rows capture the generated communication measures. We provide two measures for each communication variable and for the outcome variable of the voting procedure, absolute values and values demeaned on the group level. The relative measures are important for the first hypothesis because the leader selection process depends on a comparison of the different messages within the group, not the absolute values.[4] Thus, the regression models do not incorporate a group-specific error-term because this is already accounted for by taking the differences to the mean of the other players in the group as explanatory variables. Still, the magnitude of the explanatory variables within each group will vary considerably across groups, depending on the similarity of statements within groups (Abadie et al., 2023). We capture this

[3] The original German terms are: For the fairness concept 'gleich,' 'viertel,' 'gerecht,' 'aufteilen,' 'verteilen,' 'anteil,' 'teilen,' 'gleichmäßig,' 'teil,' 'verteilung.' For the efficiency concept '100,' '160,' 'einskommasechs,' '640,' 'gruppenkonto,' 'gruppeneinkommen.' For the modality/conditionality concept 'wenn,' 'würde.'

[4] We denote these relative variables by the prefix "My...". These variables are generated by taking the individual i's difference to the mean of the other three players in i's group. Thus, it is the deviation of a variable compared to the average of the other three group members.

variation of the treatment-assignment probabilities with standard errors clustered at the group level in all models.

Table 1: Summary Statistics

Variable	Mean	SD	Min	Median	Max
Behavioral Variables					
MyRelativeRank	3	1.7	0	3	6
TotalFollowerContributions	202	76	0	202	300
LeaderContribution	84	25	0	100	100
TotalContributions	285	94	20	300	400
FairI	0.59	0.49	0	1	1
FairII	0.75	0.43	0	1	1
Communication Variables					
Explicitness	0.67	0.24	0.09	0.74	0.97
MyExplicitness	0	0.27	-0.75	0.05	0.58
CountCharacters	320	194	4	280	854
MyCountCharacters	0	221	-458	-16	659
CountFairness	1.8	1.4	0	2	7
MyCountFairness	0	1.6	-3	-0.17	5.7
CountEfficiency	1.5	1.8	0	1	8
MyCountEfficiency	0	2.1	-4	-0.33	8
CountModals	1.5	1.4	0	1	7
MyCountModals	0	1.7	-3.3	-0.33	6

Hypothesis 1: Voting We start our analysis with Hypothesis 1 regarding the impact of the statements on the votes received (i.e., the relative rank within the group). Explicitness, as determined by the topic-model, has a positive effect on the votes that leadership candidates receive from the other group members, see model 1 in Table 2. The results in model 2 of Table 2 show that writing more about fairness or efficiency, as measured by the word count variables, attracts votes as well. The insignificant interaction term shows that addressing both topics has no complementary effect that attracts more votes. When we include both approaches in one estimation, and control for the length of statements (variable "MyCountCharacters"), see models (3) and (4), only the effects of the word count variables remain stable. Further, length itself has a positive effect on the number of votes received, possibly because followers take it as a proxy for effort or because it correlates positively with other content that followers like. In model (5) in Tables 2 and 3, we include a measure of modality, which we will need later for the comparison to leaders' behavior.

Table 2: Voting (Hypothesis 1)

	MyRelativeRank				
	(1)	(2)	(3)	(4)	(5)
MyExplicitness	0.980		0.371	0.652	0.646
	(0.420)		(0.379)	(0.386)	(0.390)
	$p = 0.021$		$p = 0.329$	$p = 0.093$	$p = 0.100$
MyCountEfficiency		0.312	0.298	0.216	0.216
		(0.056)	(0.059)	(0.059)	(0.062)
		$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.001$
MyCountFairnessConcept		0.338	0.340	0.235	0.237
		(0.078)	(0.079)	(0.093)	(0.092)
		$p < 0.001$	$p < 0.001$	$p = 0.012$	$p = 0.011$
MyCountEfficiency		−0.010	−0.015	−0.012	−0.009
×MyCountFairnessConcept		(0.020)	(0.022)	(0.020)	(0.022)
		$p = 0.613$	$p = 0.492$	$p = 0.560$	$p = 0.684$
MyCountCharacters				0.002	0.002
				(0.001)	(0.001)
				$p = 0.018$	$p = 0.051$
MyCountModals					0.008
					(0.111)
					$p = 0.940$
MyCountFairnessConcept					−0.017
×MyCountModals					(0.042)
					$p = 0.685$
Constant	3.000	2.998	2.997	2.998	3.012
	(0.000)	(0.003)	(0.004)	(0.004)	(0.035)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
No. of Observations	144	144	144	144	144
No. of Groups	36	36	36	36	36
R^2	0.025	0.250	0.253	0.297	0.297

Note: OLS estimates with standard errors clustered at the group level in parentheses.

Hence, we find support for Hypothesis 1. We can partly explain the votes a leadership candidate receives by their statement's explicitness about their planned behavior as a leader, and by how much they talk about efficiency and fairness in their statement, all relative to the other candidates' statements. More generally, we can see that the content of such statements of intent affects the choice of other players considerably.

Hypothesis 2: Follower Contribution We now look at the impact of the leaders' initial statements on followers' contributions to the public good. The coefficient for explicitness is again positive but statistically insignificant. Longer statements as well as

Table 3: Follower Contribution (Hypothesis 2)

	AverageFollowerContributions					
	(1)	(2)	(3)	(4)	(5)	(6)
Explicitness	11.995		1.581	5.984	6.113	6.186
	(8.862)		(9.872)	(10.113)	(10.227)	(6.785)
	$p = 0.178$		$p = 0.873$	$p = 0.555$	$p = 0.551$	$p = 0.364$
CountEfficiency		4.699	4.629	3.314	3.250	−0.638
		(1.992)	(2.051)	(2.349)	(2.303)	(1.728)
		$p = 0.020$	$p = 0.026$	$p = 0.160$	$p = 0.160$	$p = 0.712$
CountFairnessConcept		1.861	1.883	0.386	0.831	−1.369
		(1.962)	(1.988)	(2.289)	(3.716)	(2.462)
		$p = 0.344$	$p = 0.345$	$p = 0.866$	$p = 0.823$	$p = 0.579$
CountEfficiency		−0.295	−0.294	−0.171	−0.164	0.497
×CountFairnessConcept		(0.688)	(0.691)	(0.720)	(0.714)	(0.534)
		$p = 0.669$	$p = 0.671$	$p = 0.812$	$p = 0.818$	$p = 0.354$
CountCharacters				0.025	0.025	0.015
				(0.013)	(0.015)	(0.010)
				$p = 0.067$	$p = 0.098$	$p = 0.140$
CountModals					0.470	0.234
					(3.819)	(2.349)
					$p = 0.902$	$p = 0.921$
CountFairnessConcept					−0.316	0.059
×CountModals					(1.689)	(0.969)
					$p = 0.852$	$p = 0.951$
LeaderContribution						0.517
						(0.107)
						$p < 0.001$
I(LeaderContribution						9.191
= 100)						(5.116)
						$p = 0.075$
Constant	59.109	57.644	56.647	49.999	49.334	10.908
	(7.170)	(5.617)	(9.065)	(9.415)	(10.255)	(9.714)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.263$
No. of Observations	144	144	144	144	144	144
No. of Groups	36	36	36	36	36	36
R^2	0.013	0.100	0.100	0.125	0.126	0.497

Note: OLS estimates with clustered standard errors in parentheses.

writing about efficiency positively affect the followers' contributions, while writing about fairness does not seem to lead to higher contributions by the followers, see models (2), (3) and (4) in Table 3.

Model (6) studies to which extent group members follow the contribution of the leader which they could observe. The model contains two variables that capture the impact of the leader's contribution decision. Recall that it is a dominant strategy for the leader to contribute the entire endowment. Nevertheless, about 40 percent of the leaders contribute less. The coefficient for the leader's contribution shows that a one-unit-rise in the leader's contribution increases the average contribution of a follower by about half a unit. This imperfect reciprocity is in line with the literature (Gächter and Renner (2018); Eisenkopf and Walter (2021). On top of this marginal effect, the significant dummy $I(\text{LeaderContribution} = 100)$ shows that followers significantly reward the efficient contribution or, in turn, decrease contributions if the leader deviates from the efficient contribution. These two variables raise the R^2 of the model by a lot, indicating that they matter enormously for the followers' decisions.

In summary, our results on follower behavior show that in their voting behavior, when followers have nothing but the statements to rely on, they react to the content of the statement, but if they know hard behavioral facts in terms of the leader's contribution, they place less importance on the statements. If the leaders do not give a good example in their actions, the statement content does not help fostering higher contributions by the followers.

Hypothesis 3: Leader's fairness We hypothesized that players who address fairness also make fairer distributions. Our results in the first row of Table 4 show that leaders who talk more about fairness are more likely to distribute fairly, irrespective of the fairness definition. The effects are much more pronounced once we control for modality in the language, i.e. the use of tokens that may connect a promise of fairness to a condition. Leaders could, for example, promise to behave in a fair way if all followers contributed 100%. These leaders could then deviate from a fair allocation in an easily justifiable manner. By controlling for the count of modal expressions ['if/when,' 'would'] and an interaction with the fairness-token count, we can remove these leaders from the effect of the fairness-token count. Consequently, the effects of talking about fairness are larger (compare models (2) and (5) to models (1) and (4) in Table 4, respectively). Leaders who talk about fairness without using modal expressions are more likely to be fair.

Table 4: Distribution Decisions Leader (Hypothesis 3)

	FairI			FairII		
	(1)	(2)	(3)	(4)	(5)	(6)
CountFairnessConcept	0.055	0.095	0.082	0.023	0.071	0.065
	(0.028)	(0.039)	(0.038)	(0.025)	(0.033)	(0.029)
	$p = 0.048$	$p = 0.016$	$p = 0.031$	$p = 0.359$	$p = 0.031$	$p = 0.029$
CountModals		0.042	0.030		0.085	0.078
		(0.056)	(0.060)		(0.053)	(0.053)
		$p = 0.460$	$p = 0.618$		$p = 0.110$	$p = 0.140$
CountFairnessConcept ×CountModals		−0.026	−0.023		−0.037	−0.039
		(0.025)	(0.024)		(0.023)	(0.023)
		$p = 0.286$	$p = 0.345$		$p = 0.114$	$p = 0.094$
CountEfficiency			−0.004			−0.014
			(0.019)			(0.019)
			$p = 0.820$			$p = 0.472$
Explicitness			−0.053			−0.102
			(0.150)			(0.151)
			$p = 0.726$			$p = 0.501$
CountCharacters			0.00002			0.0002
			(0.0002)			(0.0002)
			$p = 0.936$			$p = 0.353$
LeaderContribution			0.006			0.006
			(0.002)			(0.002)
			$p = 0.009$			$p = 0.015$
I(LeaderContribution = 100)			−0.113			−0.277
			(0.162)			(0.105)
			$p = 0.485$			$p = 0.009$
Constant	0.494	0.444	0.041	0.709	0.611	0.372
	(0.067)	(0.083)	(0.160)	(0.048)	(0.070)	(0.195)
	$p < 0.001$	$p < 0.001$	$p = 0.800$	$p < 0.001$	$p < 0.001$	$p = 0.059$
No. of Observations	144	144	144	144	144	144
No. of Groups	36	36	36	36	36	36
R^2	0.023	0.032	0.085	0.005	0.029	0.074

Note: OLS estimates of linear probability models with clustered standard errors in parentheses. The dependent variables are two measures of fair behavior: FairI $\equiv I(\frac{T_L}{T_L+3T_F} \leq \frac{c_L}{c_L+c_{F_{max}}+c_{F_{med}}+c_{F_{min}}})$ and FairII $\equiv I(T_F \geq \frac{c_{F_{max}}+c_{F_{med}}+c_{F_{min}}}{3})$ with transfers T_i and contributions c_i to and from the leader L and followers F. FairI defines a leader as fair if they take a share of the public good that is smaller than or equal to their share of the contributions. FairII defines a leader as fair if they transfer at least as much to each follower as they have contributed on average. See Section 4 for more details.

Interestingly, followers do not consider modal language when making their decisions. The effect of the fairness count on voting behavior and follower contributions remains almost unchanged when including the modality variable and its interaction with the fairness count the previous estimations. See model (5) in Table 2 and model (5) in Table

3. Thus, followers do not necessarily react to the same aspects of the statements that actually predict behavior by the leaders.

6 Conclusion

We study how leadership can provide a fair and efficient solution to social dilemmas by analyzing how short written statements by leadership candidates affect leadership selection, follower contributions to a public good and predict leadership behavior. In our experiment, leadership candidates who promise to behave in a fair way in their initial statements are more likely to do so, in particular if their written promises come without modal language in the statement. However, followers respond to other verbal clues. When electing a leader, followers reward longer statements and those that are very explicit about efficiency and fairness issues but they do not consider whether promises of fairness might be conditional by using modal language. In their contribution decision, which is the main efficiency concern, followers place less importance on the statements and instead react strongly to the leaders' first moves: leaders that want to achieve an efficient outcome must lead with contributing at the efficient level, language cannot make up for lack of action.

Given the imperfect match between clues the followers respond to and those that are actually informative there seems to be much room for improvement, both in the interest of efficiency and fairness. Candidates for a leading position might optimize their statements to maximize the chance of being elected, by describing explicitly how they think the group should approach the efficiency-problem and how they will use their power to distribute profits fairly. Voters might learn to pay attention to the informative aspects of election promises. Future research should study whether the clues we found to explain leaders' and followers' behavior maintain their predictive power when the players learn about these findings.

Method-wise, our study provides further support for the usefulness of text mining procedures to extract information from text based on a relatively small sample size in economic experiments. The combination of such machine learning methods with simple word counts based on economic hypotheses is a useful addition to the toolbox of analyzing experimental data. At the same time, our study illustrates a new challenge for the pre-registration of experimental studies in that the unpredictability of the outcomes of unsupervised algorithms makes it difficult to pre-specify exact analysis plans.

A Instructions

This section contains the experimental instructions which include screenshots from the experiment. Both are translated from the original German to English.

Instructions

Welcome to this experiment!

Today you will participate in a decision experiment. You will be compensated for participating in this experiment. The amount you receive depends on your decisions and the decisions of other participants. It is therefore important that you read the instructions on the following pages carefully.

These instructions are identical for all participants.

You are not allowed to communicate with other participants for the entire duration of the experiment. We therefore ask you not to talk to each other. Violation of this rule will result in exclusion from the experiment and payment.

If there is something you do not understand, please refer back to these instructions or give us a hand signal. We will then come to you and answer your question personally.

During the experiment we do not talk about euros, but about points. So your total income will first be calculated in points. Your points will then be converted into euros at the end of the experiment, using the following conversion rate:

16 points = 1 euro

At the end of today's experiment, you will be paid in cash the points you have achieved from the experiment converted into euros. In addition, you will receive 10 euros for your punctual appearance to the experiment.

The payment procedure is organized in such a way that no participant will see what amount the other participants have received. On the following pages we will explain the exact procedure of the experiment. First, you will get an overview of the basic procedure. Then we will describe the sequence of decisions in detail. We will then show you two more examples. After that, the experiment begins.

Overview of the Experimental Procedure

At the beginning of the experiment, you will be randomly assigned to a group with three other participants. In the experiment, you make decisions within this group of a total of four group members (A, B, C and D). This is a sequence of different decisions that you make exactly once each.

The composition of your group remains the same throughout the experiment. Neither you nor the other persons learn anything about the identity of the participants in the groups – neither before nor after the experiment.

You start the experiment with a private account containing 100 points. Each member of the group decides individually how much of the 100 points they will contribute to a shared group account. The points you contribute will be deducted from your private account.

The contributed points of all four group members are added and multiplied by 1.6. The amount that comes out is called the "group income."

In a group there is one member in the position X and three members in the position Y.

The positions X and Y differ in their possibilities of action. The member in position X makes the contribution decision first. The three members in position Y see what X has contributed and then make their contribution decisions. Then the member in position X allocates a sum to herself or himself from the group income. The rest is divided equally among the other three group members.

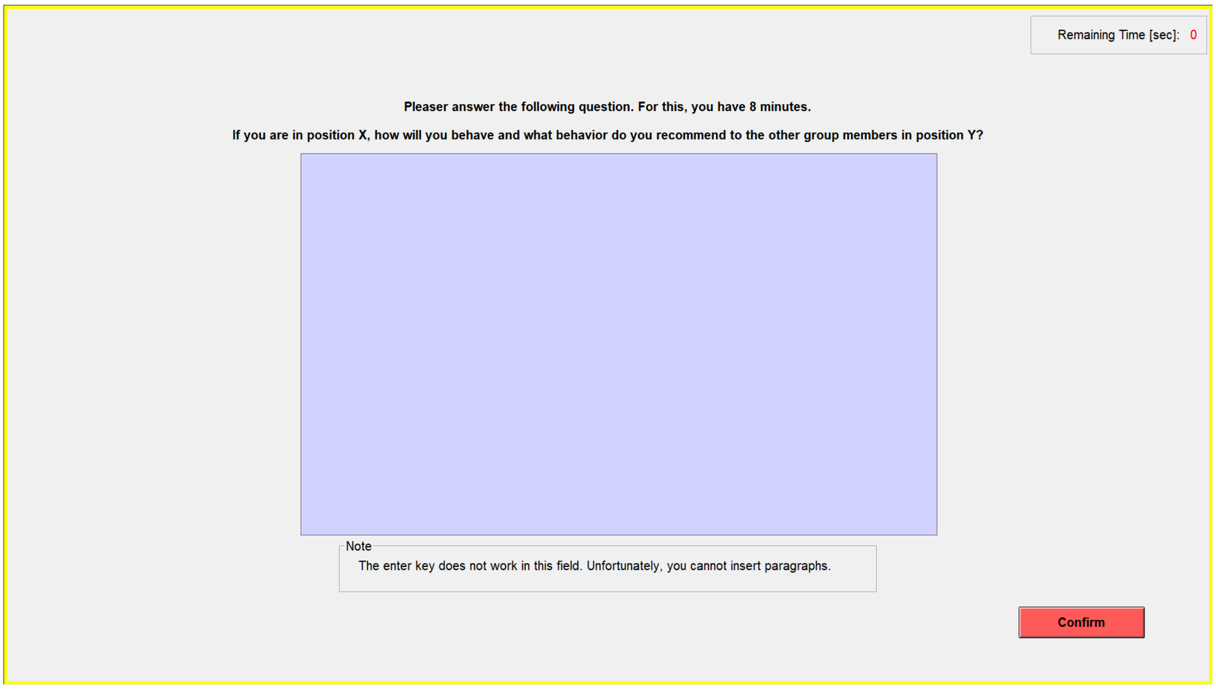
The group determines who has which position through a selection process, which we will explain in more detail on the next page.

We will explain the experiment in detail on the following pages.

Details on the Sequence of Decisions

Step 1: Message to the other group members

In the first step, you formulate a message to the other members of your group. In this message you describe how you will behave in position X.



In step 2, your message will be shown to the other group members and you will also be shown the messages of the other group members.

Step 2: Selecting the group member in position X

After you have read the messages of all group members, you enter a ranking order on the screen, indicating which member should be in position X. To do this, you and the other members of your group distribute 6 votes to each of the four group members, although each of you can also distribute votes to yourself. When distributing your votes, a "draw" is not allowed: You have to give 3 votes to one group member, 2 votes to one group member, 1 vote to one group member and 0 votes to one group member.

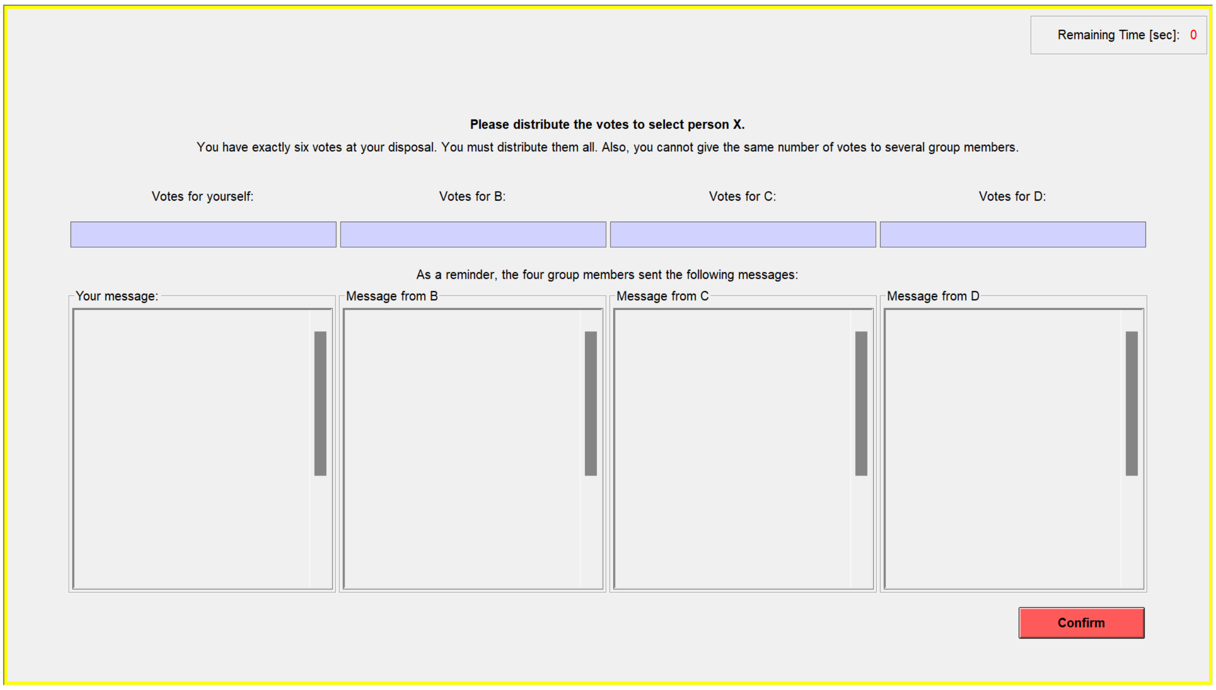
The actual selection process then works as follows:

1. First, a random mechanism of the computer program determines a group member to be the decision-person. Only the votes of this person count for the selection of

position X. Each group member has the same probability of becoming the decision-person.

2. Then a second random mechanism determines which group member is in position X – based on the votes of the decision-person. Here, the probabilities of the four group members to be in position X differ based on the distribution of votes given by the decision-person:

   - The group member to whom the decision-person gave 3 votes has the highest probability (50% or 3/6) of getting position X.
   - The group member to whom the decision-person gave 2 votes has the second highest probability (33% or 2/6) of getting position X.
   - The group member to whom the decision-person gave 1 vote has the third highest probability (17% or 1/6) of getting position X.
   - The group member with 0 votes cannot get position X.



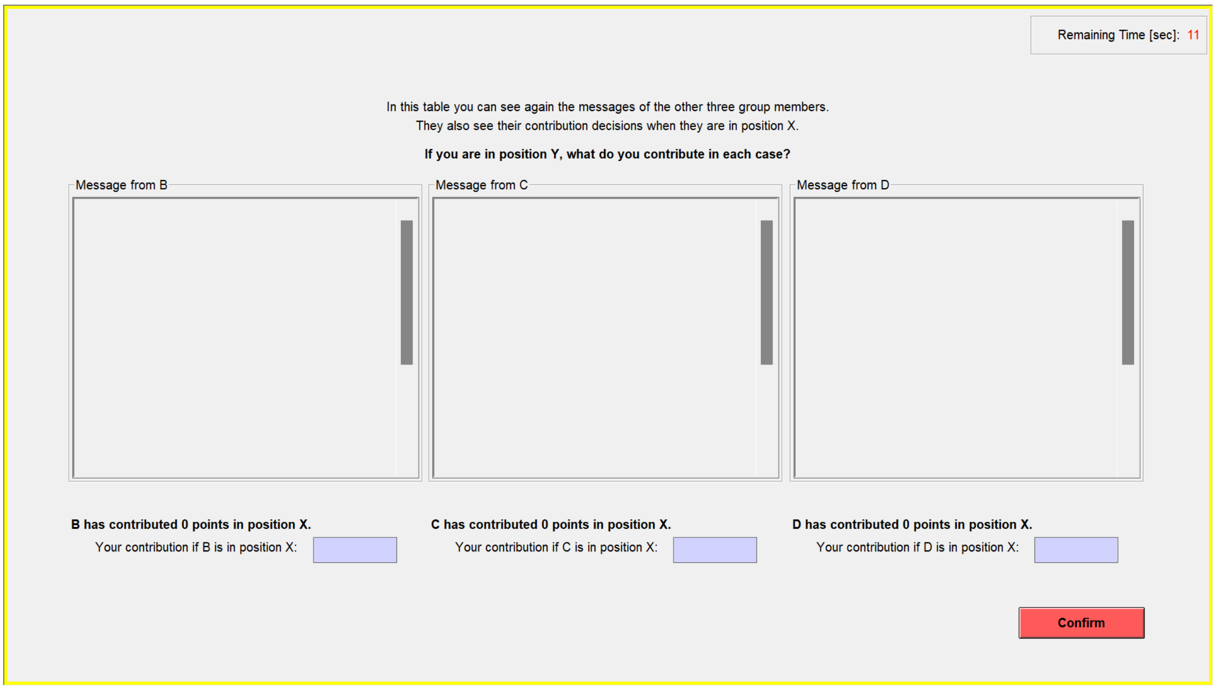
Although the selection process takes place immediately after the votes are cast, you and the other members of the group are not immediately told the result. You will only be informed at the end of the experiment which group member got into position X through this selection procedure. This means that each group member makes both the decisions in position X, and the decisions in position Y. Only at the end of the experiment you will find out which decisions are actually implemented and thus are relevant for your payout.

Step 3: Your contribution when you are in position X

Each group member decides what they will contribute to the group account when they are in position X.

Step 4: Your contribution when you are in position Y

You will see again the messages of the other three group members, and now also the contributions to the group account that the respective group member has decided to make for the case in which she/he is in position X. For each possible group member in position X, you decide on your own contribution to the group account for the case in which you are in position Y.

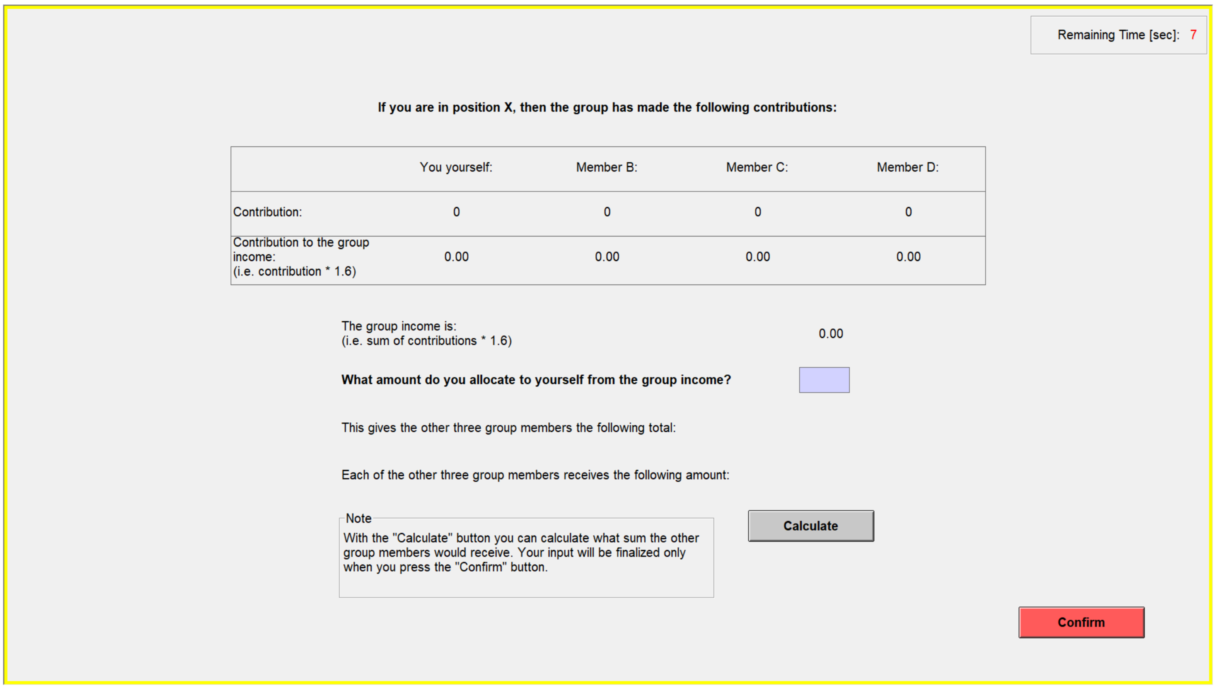


Step 5: Distribution of the group income

Now you decide how the group income will be distributed if you are in position X yourself.

To do this, the computer program first calculates the group income in case you are in position X yourself. To do this, your contribution, which you decided on in step 3, is added up with the contributions of the other three group members in step 4. This sum of the contributions of the four group members (the group account) is multiplied by 1.6. So the group income is 1.6 times the amount in the group account.

You decide which sum of the group income you will receive yourself. The rest of the group income is divided equally among the other three group members.



Step 6: Results

You will learn which group member was actually in position X, and what contributions the four group members entered in that case. You will also learn what the group income was, what amount the group member in position X allocated to herself or himself, and what amount the three group members in position Y received. You will also learn what the payoffs in points are for all four group members.

After the experiment we will ask you to fill out a short questionnaire on the computer. After that you will receive your payoff.

Examples

Example 1: The group member in position X contributes 100 points. The other three group members contribute 20, 50 and 80 points. Thus, the total in the group account is $100 + 20 + 50 + 80 = 250$ points.

The group income is then $250 * 1.6 = 400$ points.

The group member in position X allocates 400 points to himself. This leaves 0 points for the other three group members.

Thus, the income of the group member in position X is:

100 [initial endowment] - 100 [contribution] + 400 [share from group income] = 400.

The incomes of the other three group members are:

100 [initial endowment] - 20 [contribution] + 0 [share from group income] = 80

100 [initial endowment] - 50 [contribution] + 0 [share from group income] = 50

100 [initial endowment] - 80 [contribution] + 0 [share from group income] = 20

Example 2: The group member in position X contributes 0 points. The other three group members contribute 20, 50 and 80 points.

Thus, the total in the group account is $0 + 20 + 50 + 80 = 150$ points.

The group income is then $150 * 1.6 = 240$ points.

The group member in position X allocates 60 points to himself. The remaining 180 points are automatically divided equally among the remaining three group members. Thus, they each receive 60 points.

Thus, the income of the group member in position X is:

100 [initial endowment] - 0 [contribution] + 60 [share from group income] = 160.

The incomes of the other three group members are:

100 [initial endowment] - 20 [contribution] + 60 [share from group income] = 140

100 [initial endowment] - 50 [contribution] + 60 [share from group income] = 110

100 [initial endowment] - 80 [contribution] + 60 [share from group income] = 80

Now please turn to the screen. We will ask you to answer some quiz questions. This is to make sure that all participants have understood the instructions well.

If something is unclear to you, please give a clear hand signal. We will then come to your seat.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics 138*(1), 1–35.

Andres, M., L. Bruttel, and J. Friedrichsen (2023). How communication makes the difference between a cartel and tacit collusion: A machine learning approach. *European Economic Review 152*, 104331.

Arbak, E. and M.-C. Villeval (2013). Voluntary leadership: motivation and influence. *Social Choice and Welfare*, 1–28.

Bellemare, C., A. Sebald, and S. Suetens (2019). Guilt aversion in economics and psychology. *Journal of Economic Psychology 73*, 52–59.

Ben-Ner, A., L. Putterman, and T. Ren (2011). Lavish returns on cheap talk: Two-way communication in trust games. *The Journal of Socio-Economics 40*(1), 1–13.

Brandts, J., D. J. Cooper, and R. A. Weber (2015). Legitimacy, communication, and leadership in the turnaround game. *Management Science 61*(11), 2627–2645.

Bruttel, L. and U. Fischbacher (2013). Taking the initiative. what characterizes leaders? *European Economic Review 64*, 147–168.

Cappelen, A. W., B.-A. Reme, E. Ø. Sørensen, and B. Tungodden (2016). Leadership and incentives. *Management Science 62*(7), 1944–1953.

Caselli, F. and T. Cunningham (2009). Leader behaviour and the natural resource curse. *Oxford economic papers 61*(4), 628–650.

Centorrino, S. and L. Concina (2013). A competitive approach to leadership in public good games.

Dannenberg, A. (2015). Leading by example versus leading by words in voluntary contribution experiments. *Social Choice and Welfare 44*(1), 71–85.

Drouvelis, M., D. Nosenzo, and M. Sefton (2017). Team incentives and leadership. *Journal of Economic Psychology 62*, 173–185.

Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior 47*(2), 268–298.

Eichenseer, M. (2019). Leading by example in public goods experiments: What do we know? *Available at SSRN 3441638*.

Eisenkopf, G. (2020). Words and deeds – experimental evidence on leading-by-example. *The Leadership Quarterly 31*, 101383.

Eisenkopf, G. and A. Bächtiger (2013). Mediation and conflict prevention. *Journal of Conflict Resolution 57*(4), 570–597.

Eisenkopf, G. and T. Kölpin (2021). Leading-by-example: A meta-analysis. Report, University of Vechta - Unpublished Manuscript.

Eisenkopf, G. and C. Walter (2021). Leadership with imperfect monitoring. *The Leadership Quarterly*, 101589.

Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and economic behavior 54*(2), 293–315.

Feltovich, N. and P. J. Grossman (2015). How does the effect of pre-play suggestions vary with group size? experimental evidence from a threshold public-good game. *European Economic Review 79*, 263–280.

Fiala, L. and S. Suetens (2017). Transparency and cooperation in repeated dilemma games: a meta study. *Experimental economics 20*(4), 755–771.

Fiedler, M. and E. Haruvy (2009). The lab versus the virtual lab and virtual field—an experimental investigation of trust games with communication. *Journal of Economic Behavior and Organization 72*(2), 716–724.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review 100*(1), 541–556.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters 71*(3), 397–404.

Gächter, S. and E. Renner (2018). Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior and Organization 154*, 321–334.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association 1*(1), 114–125.

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

Haigner, S. D. and F. Wakolbinger (2010). To lead or not to lead: Endogenous sequencing in public goods games. *Economics Letters 108*(1), 93–95.

Karakostas, A., M. G. Kocher, D. Matzat, H. A. Rau, and G. Riewe (2023). The team allocator game: Allocation power in public goods games. *Games and Economic Behavior 140*, 73–87.

Koukoumelis, A., M. V. Levati, and J. Weisser (2012). Leading by words: A voluntary contribution experiment with one-way communication. *Journal of Economic Behavior and Organization 81*(2), 379–390.

Ledyard, J. O. (1995). *Public Goods: a Survey of Experimental Research*, pp. 111. Princeton: Princeton University Press.

Nielsen, K., P. Bhattacharya, J. H. Kagel, and A. Sengupta (2019). Teams promise but do not deliver. *Games and Economic Behavior 117*, 420–432.

Pogrebna, G., D. H. Krantz, C. Schade, and C. Keser (2011). Words versus actions as a means to influence cooperation in social dilemma situations. *Theory and Decision 71*(4), 473–502.

Préget, R., P. Nguyen-Van, and M. Willinger (2016). Who are the voluntary leaders? experimental evidence from a sequential contribution game. *Theory and Decision 81*(4), 581–599.

Rivas, M. F. and M. Sutter (2011). The benefits of voluntary leadership in experimental public goods games. *Economics Letters 112*(2), 176–178.

Rosenbaum, S. M., S. Billinger, and N. Stieglitz (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology 45*, 181–196.

Ross, M. L. (2015). What have we learned about the resource curse? *Annual review of political science 18*, 239–259.

Serra-Garcia, M., E. Van Damme, and J. Potters (2011). Hiding an inconvenient truth: Lies and vagueness. *Games and Economic Behavior 73*(1), 244–261.

Sheremeta, R. M. and J. Zhang (2014). Three-player trust game with insider communication. *Economic Inquiry 52*(2), 576–591.

Song, F. (2008). Trust and reciprocity behavior and behavioral forecasts: Individuals versus group-representatives. *Games and economic behavior 62*(2), 675–696.

Spadaro, G., C. Graf, S. Jin, S. Arai, Y. Inoue, E. Lieberman, M. I. Rinderu, M. Yuan, C. J. Van Lissa, and D. Balliet (2022). Cross-cultural variation in cooperation: A meta-analysis. *Journal of Personality and Social Psychology*.

Stoddard, B. V., C. A. Cox, and J. M. Walker (2021). Incentivizing provision of collective goods: Allocation rules. *Southern Economic Journal 87*(4), 1345–1365.

Van der Heijden, E., J. Potters, and M. Sefton (2009). Hierarchy and opportunism in teams. *Journal of Economic Behavior & Organization 69*(1), 39–50.

Vanberg, C. (2008). Why do people keep their promises? an experimental test of two explanations1. *Econometrica 76*(6), 1467–1480.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics 6*(3), 299–310.