Communication in a threshold public goods game with ambiguity

Matteo M. Marini[a], Aurora García-Gallego[b], and Luca Corazzini[c]

[a]Department of Economics and Management, University of Florence, Florence, Italy

[b]LEE and Department of Economics, Universitat Jaume I, Castellón, Spain

[c]Department of Economics, University of Venice "Cà Foscari", Venice, Italy

Abstract

This paper offers evidence on the impact of communication on the provision of public goods whose quality is uncertain, such as investment in prevention or fundraising in favor of non-governmental organizations (NGO). We run a laboratory experiment with two treatments, where the control variable is pre-play communication via unrestricted text chat. A binary threshold public goods game with four-person groups, threshold of three contributors and provision mechanism with ambiguity is at the core of the design. A private signal linked to the real value of the public good supports the contribution decision. We

This paper is based on a chapter from Matteo M. Marini's PhD thesis.

Correspondence to: Department of Economics and Management, Faculty of Economics, University of Florence, Via delle Pandette 9, 50127 Florence (Italy).

E-mail addresses: matteom.marini@gmail.com (M.M. Marini), mgarcia@eco.uji.es (A. García-Gallego), luca.corazzini@unive.it (L. Corazzini).

Declarations of interest: none.

find that unexpected low-valued public goods can undermine future willingness to contribute, although the benefits of communication still prevail in the form of higher public good provision. We also detect unprecedented inefficiency coming from overcontribution, given that subjects tend to neglect the free-rider problem. Chat analysis reveals that players favor the minimization of ambiguity over the maximization of the group earnings, so that we finally speculate that under uncertainty satisficing is more salient than optimizing.

Keywords: communication, ambiguity, experiment, private signal, threshold public goods game.
JEL classification codes: C92, D81, D82, D83, H41.

1 Introduction

Decision makers often end up with private information as the only estimate of common-value public goods whose profitability is pervaded by uncertainty (Cox, 2015). For instance, investment in prevention typically produces variable outcomes due to the unpredictable nature of natural disasters and epidemics, in such a way that even the investors themselves cast doubt on the efficacy of such actions. In these environments, should the members of a community be able to aggregate information, uncertainty would definitely be reduced.

As motivating real-world situation where the quality of the final output cannot be inferred in advance, here we consider fundraising in favor of a non-governmental organization (NGO) that aims at constructing a new school in a developing country. Suppose this NGO arranges a donation campaign on its website such that donors are allowed to give only a preset amount (e.g., 5-dollar donations), as well as a threshold sum of money is needed to accomplish the goal. Moreover, the developing country in question is currently subject to war and political instability to the extent that, even in case the targeted money is raised, the school might never be built. Unless the NGO already has a reputation, uncertainty is also related to how effectively this scenario is handled. However, in order to streamline fundraising the NGO involves well-established communities in the humanitarian sector (e.g., e-clubs associated with Rotary International) whose members hold some private information about the chances that the school will actually be constructed. For instance, a donor might have a deep knowledge either of the country or of the NGO's modus operandi such that its disclosure might be beneficial from the perspective of group decision making, thereby reducing uncertainty. At the same time, information aggregation might raise perceived quality and expectations insofar as the occurrence of an unsuccessful social project leaves donors disappointed, impairing future willingness to donate. In other words, unexpected low quality of the public good could hamper successive contributions. In favour of this argument, disappointment originates from the comparison between a given event and a better state of the world which has not materialized, taking the shape of "a psychological reaction to an outcome that does not match up to expectations"(Bell, 1985). Furthermore, there is experimental evidence that people are particularly "sensitive to negative consequences that occur as a result of their own actions" (Böhm et al., 2016), this begetting a behavior change in the subsequent interaction. From the viewpoint of efficiency, making a donation without reaching the monetary goal is definitely wasteful, but exceeding the threshold may reveal minor inefficiency as well, whenever the extra money cannot significantly improve the quality of the public good.

Accordingly, this paper aims to test whether the opportunity to communicate actually yields efficient results also under stochastic values of the public good, with a particular focus on the consequences of low-quality outcomes. To the best of our knowledge, this is the first study investigating cooperation under uncertain values of threshold public goods (TPG), in that scholarly efforts to date have concentrated on linear public goods (Gangadharan and Nemes, 2009; Artinger et al., 2012; Cherry et al., 2015). This allows us to analyze the interplay between the issue of ambiguity and the coordination problem, especially in the presence of a trade-off. By so doing, we also answer Bardsley and Ule (2017)'s call for more attention on team reasoning, which the authors regard as "a key ingredient of the explanation of coordination game data". Our study tests this theory in a context where the predictions thereof are in conflict with the minimization of uncertainty.

As to the experimental setup, we equally split the participants into two treatments, the control variable being pre-play communication in the form of unrestricted text chat. A binary TPG game with four-person groups and a threshold of three contributors is at the core of the design, which also includes secondary tasks controlling for altruism and risk preferences, among other possible moderators. Uncertainty typifying the donation process is rendered through a novel provision mechanism, that is, a binary lottery producing either a high or a low public good value for the whole group, with constant unknown probability across rounds. An individual binary signal linked to the lottery outcome is private knowledge and supports subjects' contribution decision.

The results emphasize that unexpected low-valued public goods can actually undermine future willingness to contribute only when communication is allowed. In any case, the benefits of communication prevail over the drawbacks thereof, since coordination leads to higher public good provision and reduces wasteful undercontribution in the chat treatment. We also detect unprecedented inefficiency in the form of massive overcontribution, given that subjects tend to neglect the free-rider problem. Chat analysis reveals that the group agreements on all members contributing are often successful, arguably thanks to pronounced group identity as well as players' willingness to pursue symmetric payoffs.

To conclude, despite the presence of a valid coordination device as the rotation of the non-contributor, the group members prefer the implementation of strategies based on the systematic disclosure of all four private signals, or in other words, the minimization of uncertainty is favored over the maximization of the group earnings.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 includes a detailed description of the experimental phases as well as the hypotheses of this work. The results are presented in Section 4, where after performing an overall analysis, we test such hypotheses and propose possible explanations. Section 5 concludes.

2 Literature review

Prominent surveys on public goods games (Ledyard, 1995; Chaudhuri, 2011; Dhami, 2016) identify communication as one of the non-punitive methods in order to sustain cooperation. Although face-to-face communication turns out to be the most effective form, still communication via computer monitor outperforms the no-communication condition on promoting cooperation (Bochet et al., 2006). After reviewing the experimental literature on social dilemmas, Bicchieri and Lev-On (2007) theorize the positive influence of the communication effect also in computer-mediated environments, by stressing that the effect is as powerful as these environments are able to reproduce the features of face-to-face communication. Palfrey et al. (2017) add that the effect of pre-play communication on efficiency depends on the richness of the message space, since the efficient outcomes observed under unrestricted text chat are not found when binary or numerical messages are employed. From a theoretical viewpoint, Agastya et al. (2007) indicate that binary communication can provide individuals with incentives to contribute, as compared to the absence of communication. As far as coordination games are concerned, the picture becomes more complicated: one-way communication should be preferred in games of conflict such as the battle of the sexes, whereas two-way communication is more effective when there is no dominated strategy (Cooper et al., 1992). Non-binding pre-play communication does not always entail efficiency gains, which still occur in case players decide to disclose private information (Crawford and Sobel, 1982). By and large, a number of moderators whereby communication operates have been detected, such as the development of group identity, the change of expectations of others' behavior and the offering of promises (Ostrom et al., 1994). Nevertheless, even deviations from pledges on Nash equilibria are observed (Tavoni et al., 2011).

The current work approaches the topic within the framework of TPG games, where the public good value is provided as long as a certain threshold is overcome. Also known as step-level public goods or public goods with a provision point mechanism, common examples are the construction of infrastructures, the passing of a law in parliament and fundraising projects, whose quality or quantity do not increase significantly if the contributions exceed the threshold (Offerman et al., 1996). As Alberti and Cartwright (2016) point out, a relevant difference from standard public goods is that there is not the typical conflict between the maximum individual payoff and the best social outcome, since some Nash equilibria coincide with the most efficient combination of strategies. This implies that, in addition to cooperation, the current problem becomes a matter of coordination. In the last decade, TPG games have been repeatedly employed in the context of environmental issues, notably climate change, where each individual is supposed to make an effort so as not to surpass thresholds of environmental resources whose excessive use could lead to harmful consequences (Tavoni et al., 2011; Barrett and Dannenberg, 2012; Güth et al., 2015). In this frame, elements of ambiguity are often introduced to simulate uncertainty with respect to the level of the threshold, which usually is unknown.

Ambiguity (i.e., uncertainty) refers to the case where the outcome probabilities of a risky choice are not known by the decision makers. On the one hand, the subject has been studied long in the field of individual decision making after Ellsberg (1961)'s criticism of Savage's expected utility theory with subjective probabilities, mainly by postulating individual ambiguity aversion. On the other hand, the interest in delving into ambiguity within strategic environments, where the concept is generally elaborated as uncertainty about the behavior of the others, has arisen only recently. Using this definition, there is evidence that ambiguity is able to blunt the free-riding problem (Eichberger and Kelsey, 2002; Bailey et al., 2005; Keenan et al., 2006), as well as that safe strategies appear more attractive under uncertainty, which is consistent with the ambiguity aversion hypothesis (Kelsey and Le Roux, 2017; Calford, 2017).

3 Experimental design and hypotheses

The experiment is programmed in z-Tree (Fischbacher, 2007) and takes place at the Laboratorio de Economía Experimental (LEE) of Universitat Jaume I (Castellón, Spain). Overall, we recruite 160 students (73 females) from the university campus through ORSEE (Greiner, 2015) and equally split them into two treatments.

As illustrated in Table 1 and Table 2, a TPG game is at the core of the design, which also encompasses a Ring-test capturing social orientation and a few questionnaires measuring variables of interest.[1]

Table 1: Structure of the experiment

Treatment	N° of sessions	Subjects	Females
No communication (NC)	2	80	50%
Chat	2	80	41%

The treatment variable is contained in the TPG game and represented by unbinding pre-play communication in the form of unrestricted text chat. Throughout the experiment salient financial incentives are provided, since one period of the TPG game is randomly chosen and a further lottery between that round and the Ring-test outcome determines the final payment, in addition to a show-up fee of 5 euro.[2] The control treatment takes one hour and a half to be carried out. We increase the show-up fee to 7 euro in the chat sessions, which last two hours. The experimental currency is the ECU (Experimental Currency Unit), with the exchange rate 5 ECU = 1 euro in both the remunerated tasks of the experiment. In the end, the subjects in the baseline gain on average 18.4 euro, whereas the mean earnings in the chat treatment amount to 20.5 euro.

3.1 Threshold public goods game

In each session, after the roll call the subjects are randomly allocated to the cubicles and individually read the instructions of the first task. Before starting the TPG

[1]As to the rationale behind the order of phases, incentivized tasks have priority over questionnaires. We decide to begin sessions with the TPG game given that the research question is based on it. While the socio-demographic queries are assumed to be neutral and placed between tasks, interdependency and empathy questions are answered at the end since they would probably influence subsequent behavior in compliance with the theory of cognitive dissonance.

[2]The random outcomes of the two lotteries vary from subject to subject.

Table 2: Structure of each session

Order	Phase
1	Control questions
2	TPG game with(out) chat
3	Socio-demographic questionnaire
4	Ring-test
5	Interdependency and empathy questions
6	Payment feedback

game, the participants are supposed to answer five control questions by which we want to test their grasp of rules and incentives provided.[3] Once the solutions are explained aloud, the game begins and four-person groups are formed with partner matching over fifteen periods. At the beginning of every period, each group member is endowed with 50 ECU and can either contribute the entire endowment to a public good or decide not to contribute by bearing a smaller cost equal to 15 ECU.[4] In other words, this is a binary TPG game where a threshold number of contributors are needed for the production of the public good (Dawes et al., 1986; Palfrey and Rosenthal, 1991; Offerman et al., 1998).

As compared with related literature, the provision mechanism features a novel concept of ambiguity and consists in a binary lottery producing either a high value of 80 ECU or a low value of 20 ECU for each group member. The lottery is based on a box containing 100 balls, where each of them holds a 70% chance of being white.[5] This probability is concealed from the subjects, who only know that the chance is constant throughout the treatment. Every period, in case the threshold of three contributions out of four is reached, one ball is drawn: if the ball is white, each member wins 80

[3]Full experimental instructions are reported in Appendix A along with the five control questions.

[4]We introduce the cost of not contributing because the decision of not donating entails a psychological cost due to guilt aversion.

[5]This implies that the number of white balls in the box varies across rounds.

ECU. Otherwise, 20 ECU. Whenever the threshold is not achieved, the contributions are not refunded and the lottery is not played.

Prior to making the contribution decision, every player is provided with a private signal which is dichotomized for the sake of interpretation. Each group member draws an independent sample of 10 balls from the box and subsequently receives a message which can be of two kinds:

- the sample proportion of white balls is high (greater than or equal to 5/10);
- the sample proportion of white balls is low (less than 5/10).

At this point, the players in the treatment sessions are allowed to send messages to the other group members for an interval of 90 seconds. Once time expires, the chat freezes but the messages exchanged remain visible until the participants make the contribution decision.

Given a group made up of N individuals, the public good is yielded only if at least K members contribute. Accordingly, the return net of endowment for player i amounts to:

$$
\begin{array}{rccl}
-15+g & \text{if} & i & \text{does not contribute and at least } K \text{ others contribute} \\
-15 & \text{if} & i & \text{does not contribute and fewer than } K \text{ others contribute} \\
-50+g & \text{if} & i & \text{contributes and at least } K \text{ - 1 others contribute} \\
-50 & \text{if} & i & \text{contributes and fewer than } K \text{ - 1 others contribute}
\end{array}
$$

where $N = 4$, $K = 3$ and $g \in \{20, 80\}$.

Following Palfrey et al. (2017)'s modus operandi, the game involves three sources of inefficiency: (a) undercontribution, in case just one or two group members contribute and the threshold is not achieved; (b) overcontribution, when all four players contribute and the public good is provided at excessive cost; (c) the usual free-rider problem, since everyone would prefer to enjoy the public good, whether high-valued or low-valued, without contributing.

Our design also takes into account the individual belief about what the other group members do, given that everyone would like her own contribution to be critical for the threshold achievement (Offerman et al., 1996). As in Caporael et al. (1989), the single contribution can be categorized as *futile* if fewer than K - 1 others contribute, *critical* if exactly K - 1 others contribute or *redundant* if K or more others contribute. Since Offerman et al. (1996) notice a positive correlation between the subjective probability of being critical and the contribution choice, we induce the participants to truthfully reveal their own subjective probability of being critical by means of an incentivized question which is placed immediately after the contribution stage.[6] Hence, in each period the individual profit is the algebraic sum of three components: the initial endowment, the net return and the gain from the incentivized question.

The last screen of every period gives subjects information about the individual profit, the threshold achievement, and the outcome of the public good lottery in case the latter is played. A history table summarizes relevant variables of all periods and stays visible in all stages of the treatment.

3.2 Ring-test and questionnaires

At this point, through a socio-demographic questionnaire we collect data on gender and self-reported risk aversion measured on a 1-7 rating scale, where "1"corresponds to "I love taking risks"and "7"to "I avoid taking risks". Dohmen et al. (2011) con-

[6]In order to induce risk neutrality and obtain reliable answers, we combine the quadratic scoring rule with a binary lottery procedure, as described in Harrison et al. (2014). By so doing, the more accurate the prediction, the higher (lower) the chances of winning 5 extra ECU (nothing). Being θ one's subjective probability that exactly two of the other three group members contribute, the chance of winning 5 extra ECU is equal to $P(\theta \mid C) = 100 - 100(1 - \theta/100)^2$ if actually two of the other three group members contribute. Otherwise, that probability is equal to $P(\theta \mid NC) = 100 - 100(0 - \theta/100)^2$. We minimize the possible hedging problem by keeping the stakes for belief elicitation "small" relative to the other choice tasks, in compliance with the literature (Blanco et al., 2010).

firm validity of similar stated risk preferences as compared to incentive-compatible measures, so that we are going to include this variable in the regressions at a later stage.

We also remind subjects that the next task is unrelated to the first part of the experiment and we hand out the instructions of the Ring-test, also known as Decomposed Game (Liebrand, 1984). By so doing we want to control for social orientation, namely, the weight that one ascribes to her own welfare compared to the others' well-being. Indeed, there is evidence that prosocials are more collaborative than proselfs in public goods scenarios (Parks, 1994). Afterwards we explain aloud the rules and, once all doubts are clarified, the game starts.

In this instance each player is randomly matched with another participant, whose identity remains unknown throughout the task. Each subject faces 24 scenarios and in each of them she is supposed to choose between two options, "A"and "B". Each option assigns a positive or negative amount of ECU to the subject herself (payoff x) and to her partner (payoff y). In other words, the players are asked to pick their preferred money allocation between the two options displayed in each stage. Until the end of the session, no feedback is given about the partner's choices and about the individual profit, which is calculated by adding the amount the subject assigns to herself and the money she receives from the partner, in all scenarios. To avoid potential losses, this algebraic sum is in turn added to (subtracted from) an initial endowment of 40 ECU.

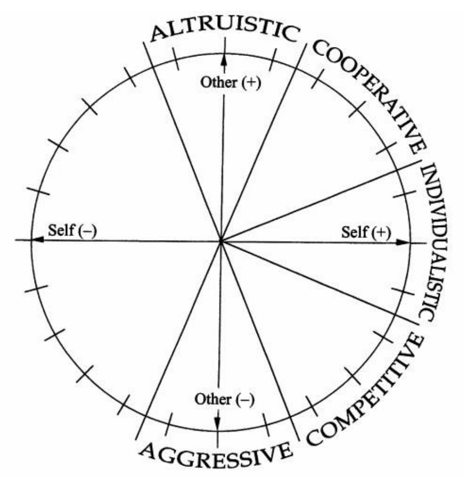
The monetary values of the payoffs are chosen so that, considering a two-dimensional space, the ordered pairs (x, y) are placed at 24 equally spaced points on the circumference of a circle, which is centered at the origin (0, 0) and has a radius of 15 monetary units. This entails that $x^2 + y^2 = 15^2$, where $x + y$ is not constant. Basically, in each scenario the player is supposed to choose between two adjacent points on this circumference.[7]

[7]We faithfully replicate the protocols in Balafoutas et al. (2013) and report the monetary values of the payoffs in Appendix B. For further details about the game, the reader is referred also to

In the end, we build a motivational vector indicating the individual's type, as in Figure 1:

- *aggressive*, with an observed vector lying between -112.5° and -67.5° (measured from the abscissa);
- *competitive*, between -67.5° and -22.5°;
- *individualistic*, between -22.5° and 22.5°;
- *cooperative*, between 22.5° and 67.5°;
- *altruistic*,[8] between 67.5° and 112.5°.

Figure 1: The value orientation circle by Offerman et al. (1996)



Offerman et al. (1996) and Brosig (2002).

[8]In Section 4 we code the observed angle as *Altruism*.

Moreover, the length of the motivational vector usually serves as indicator of consistency, that is, it shows how often the subject picked the own-other payoff combination which is the closest to her own motivational vector. If the individual chooses consistently throughout the 24 scenarios, the ensuing vector length is equal to 30. Random choices produce a vector of zero length. In our case, we decide not to drop observations given that only one subject's vector length is below 7.5, the benchmark used by Brosig (2002).

After completing the Ring-test, the participants are asked to fill in two questionnaires measuring the degree of subjective interdependency perceived during the TPG game and context-independent empathy, respectively.[9]

Finally, each subject is informed of her own wealth accumulated during the whole experiment and of the profit relevant to the final payment, which comes from a fifty-fifty lottery between the Ring-test outcome and the period randomly selected from the TPG game.

3.3 Hypotheses

In order to check if low quality of the public good has detrimental effects on the successive decision to contribute, we develop the following hypothesis at the individual level:

[9] As in Sonnemans et al. (1998), subjective interdependency is measured by means of a 7-point scale for the five following questions: (1) how much influence do you think you had on your own payoffs; (2) how much influence do you think you had on the earnings of other group members; (3) how much influence do you think you had on the decisions of others; (4) how much influence do you think the others had on your earnings; and (5) to what degree do you agree with the statement "In this experiment I and the other group members depend on each other for good results". In the results section we average out the scores from the five answers under the variable *Interdependency*.

The second questionnaire is called Empathy Quotient (EQ) and aims to measure empathy as dispositional trait consisting of both affective and cognitive components. Here participants go through 40 empathy items and 20 filler items, and can earn 2, 1, or 0 points on each empathy item so that one's final score ranges from 80 to 0. Full details are reported in Baron-Cohen and Wheelwright (2004); Lawrence et al. (2004). In the results section we code the individual score as *Empathy Quotient*.

H1. *The contribution decision is negatively affected by previous low quality of the public good.*

H1 is assumed to be true especially when the decision makers are allowed to communicate and aggregate information. Indeed, since private signals are correlated with the occurrence of high-valued public goods (on average equal to 70%), the sharing of such signals is conducive to high expectations which in case of poor quality would be disappointed. An equivalent hypothesis to H1 is also supported by Böhm et al. (2016) and justified by the so-called omission bias, according to which "people are more sensitive to negative consequences that occur as a result of their own actions than as a result of inactions". Hence, a change in future behavior is expected.

We also test four regularities related to efficiency of communication in TPG games, starting from the outcomes observed by Palfrey et al. (2017) which could be extended to our design. Here we formulate the following hypotheses at the group level:

H2. *The likelihood of public good provision is greater in the chat treatment than in the baseline.*

H3. *The average group earnings are higher in the chat treatment than in the baseline.*

H4. *The incidence of just one or two group members contributing is lower in the chat treatment than in the baseline.*

H5. *The incidence of all group members contributing is lower in the chat treatment than in the baseline.*

H2 embodies the well-documented positive effect of communication. If supported, we anticipate that this hypothesis begets H3, since the lottery associated with the public good provision yields the high value on average with 70% probability.

H4 and H5 indicate that communication improves coordination on either three or zero contributions, as theoretically predicted. Once communication is allowed, the players' beliefs about the value g of the public good should drive the outcome to either one of the two combinations, given that any agreement on four contributions is not credible due to the free-rider problem. Thus, we assume that the inefficient cases of undercontribution and overcontribution are reduced in the chat treatment.

Moreover, a valid coordination device allows to maximize the group earnings and is made available thanks to partner matching, that is, a rotation of the only one group non-contributor that solves the problem of free-riding. Such an agreement on three contributions also represents the prediction of theories of team reasoning, which claims that group identity leads each member to detect and implement a strategy achieving the maximum group payoff, as well as to play her own part within such a joint action (Bacharach, 1999; Sugden, 2000, 2003). Given its great explanatory power of coordination game data (Bardsley and Ule, 2017), we test team reasoning through chat analysis and construct H6:

H6. *In the chat treatment, rotation-based strategies are more popular than signal-based strategies.*

By signal-based strategies we mean an action profile based on the systematic sharing of all four private signals of the group, which aims at reducing uncertainty and entails decisions period by period. As a result, rotation-based and signal-based strategies are not jointly feasible, or in other words, maximizing the group payoff is in conflict with the minimization of uncertainty.

4 Results

In this section we introduce the results of the experiment by splitting them into three blocks.

First of all, we carry out an overall analysis that presents summary statistics at the individual level and checks the performance of H1. Then we test the four group-level hypotheses. The third subsection turns to the chat messages to deal with H6, finally proposing possible explanations.

4.1 Overall analysis

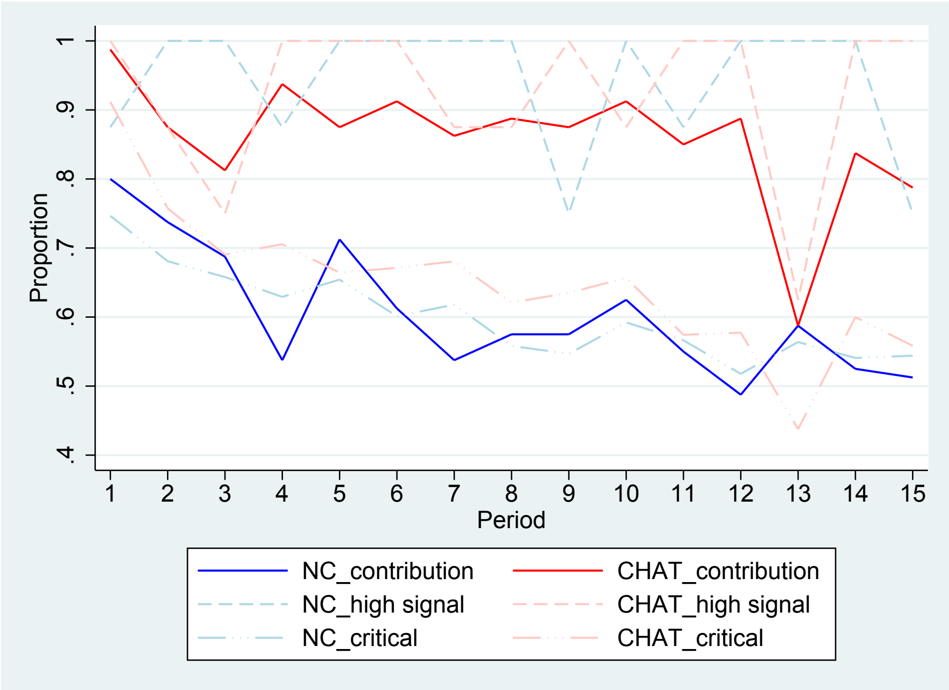
As a first step, we define the following variables:

- *contribution*, a dummy equal to 1 if player i decides to contribute, and 0 otherwise.
- *high signal*, a dummy equal to 1 if player i receives a private signal stating that the sample proportion of white balls is high, and 0 otherwise.
- *critical*, the subjective probability of being critical for the threshold achievement as revealed by player i.

Figure 2 shows how these variables change over time by comparing the no-communication (NC) condition and the chat treatment.

The main insight from the chart is that the participants in the chat sessions (M = 0.86, SD = 0.35) opt to contribute more often than the players who are not allowed to communicate (M = 0.60, SD = 0.49). Across conditions, the contribution decision seems to be driven by both high signals (M = 0.93, SD = 0.25) and subjective pivotality (M = 0.63, SD = 0.34), this being particularly evident at period 13 as far as the chat treatment is concerned. Here, the negative peaks suggest that subjects make careful use of low signals in the context of their signal-based group strategies.

Figure 2: Contributions, high signals and subjective pivotality by treatment and period



Moreover, the average Ring-test score is 15.51° (SD = 24.30), with altruistic and cooperative participants (i.e., prosocials: M = 0.75, SD = 0.43) contributing more frequently than individualistic, competitive and aggressive subjects (i.e., proselfs: M = 0.72, SD = 0.45).[10]

These preliminary remarks are supported by the random effects probit regressions in Table 3, where we tackle the issue of dependence across players by clustering standard errors at the group level. Column (1a) shows that the coefficient of *high signal* is positive and significant at the 1% level, namely, the subjects closely follow the suggestions given by their own signals. But in particular, the positive and significant coefficient ($p < 0.01$) of the treatment dummy *chat* highlights the crucial role played by communication. This effect appears meaningful, given that we also

[10]No differences in the distribution in categories are found between treatments (Fisher's exact test, $p = 0.443$). The related table is reported in Appendix B along with other supplementary data.

control for information aggregation by including a variable for the number of signals one has access to. At the same time, the contribution decision is predicted by subjective pivotality and negatively affected by participants' experience ($p < 0.01$ and $p = 0.02$, respectively), in line with related literature. Specification (2a) includes regressors from minor tasks, calling attention to the beneficial impact of prosociality on cooperation ($p = 0.015$) and the irrelevance of individual risk preferences ($p = 0.696$).

We test H1 by checking the role of the group ability to reach the threshold in the previous period but, quite surprisingly, the coefficient of the lagged *low-valued PG* is not statistically significant even after we consider the interaction with the treatment dummy. Conjecturing that partner matching and group strategies might entail persistence of players' choice, in columns (1b) and (2b) we analyze the dynamics of contribution over time and implement Blundell and Bond (1998) two-step system GMM estimators, which are robust under heteroskedasticity and let us include time-invariant regressors. As already debated but not sufficiently operationalized in the experimental literature (Brañas-Garza et al., 2011, 2013; Bortolotti et al., 2015), dynamic panel data models allow to unveil new relationships between variables as compared to static regressions, and our work is no exception. Indeed, in both the dynamic specifications the interaction term reaches statistical significance at the 5% level ($p = 0.010$ and $p = 0.021$, respectively) and emphasizes that poor quality of the public good has a negative impact only in the presence of communication. This is consistent with information aggregation raising expectations about the TPG value, as well as with the aforementioned omission bias:

Result 1. *The contribution decision is negatively affected by previous low quality of the public good (support for H1). This outcome emerges only in the chat treatment thanks to the use of dynamic panel data techniques.*

Table 3: Explaining contributions

	Dependent variable: contribution decision			
Technique	(1a) SPDM	(1b) DPDM	(2a) SPDM	(2b) DPDM
Contribution		0.078*		0.076*
(t - 1)		(0.042)		(0.042)
Chat	0.935***	0.186***	0.898***	0.160**
	(0.344)	(0.071)	(0.318)	(0.067)
Low-valued PG	-0.035	0.026	-0.040	0.017
(t - 1)	(0.162)	(0.037)	(0.163)	(0.037)
*Low-valued PG * Chat*	-0.196	-0.122**	-0.190	-0.106**
(t - 1)	(0.258)	(0.047)	(0.257)	(0.046)
Trend	-0.027**	-0.014***	-0.027**	-0.014***
	(0.012)	(0.005)	(0.012)	(0.005)
Signals accessed	0.223**	0.047**	0.217**	0.043**
	(0.097)	(0.019)	(0.097)	(0.018)
High signal	1.077***	0.184***	1.080***	0.186***
	(0.186)	(0.059)	(0.185)	(0.058)
Critical	0.022***	0.005***	0.022***	0.005***
	(0.003)	(0.001)	(0.003)	(0.001)
Female			-0.418**	-0.137**
			(0.174)	(0.053)
Risk aversion			0.023	0.004
			(0.059)	(0.016)
Altruism			0.009**	0.002*
			(0.004)	(0.001)
Empathy Quotient			-0.001	0.004*
			(0.010)	(0.003)
Interdependency			0.191	0.017
			(0.120)	(0.035)
Constant	-2.023***	0.198**	-2.784***	-0.013
	(0.307)	(0.097)	(0.873)	(0.202)
Wald χ^2	174.010***	208.570***	231.800***	251.010***
Hansen test		0.135		0.194
AB test for AR(2) in fd		0.193		0.211
No. of instruments		113		118
No. of observations	2,240	1,120	2,240	1,120

(1a) and (2a): coefficient estimates from random effects probit models, with standard errors clustered at the group level in parentheses. (1b) and (2b): coefficient estimates from the implementation of two-step system GMM estimators, with standard errors corrected à la Windmeijer (2005) in parentheses. *Contribution (t - 1)* is the contribution decision of the subject in the previous round (i.e., the only endogenous regressor of the dynamic models). *Chat* is the treatment dummy. *Trend* is a linear time trend that starts from 0. *Low-valued PG (t - 1)* is a dummy equal to 1 if subject's group reached the threshold and obtained a low-valued public good in the previous period. p-values for Hansen test (with null hypothesis that the overidentifying restrictions are valid), and for Arellano-Bond test for second-order autocorrelation in first differences.
*** p-value < 0.01.
** p-value < 0.05.
* p-value < 0.10.

In any case, the mild evidence for persistence of the subjects' decision ($p = 0.063$ and $p = 0.072$) indicates that, albeit mostly long-term, both rotation-based and signal-based strategies do not entail the same contribution decision across periods. All the other results of the static models are confirmed.

4.2 Group outcomes

Considering each group as independent observation, Table 4 compares the frequency of public good provision and the average group earnings in the two treatments by means of a one-tailed Wilcoxon rank sum test.

As the considerable discrepancy in contributions at the individual level already suggested, the overall provision rate is significantly greater ($p < 0.01$) when the subjects are free to exchange messages. The outcome is not different if we take into account any subset of periods. In the chat treatment the decline of public good provision is really subtle and the groups reach the threshold on average in 92% of cases, versus 56% in the baseline. On the contrary, the decrease of provision rate is more pronounced in the absence of communication. Therefore:

Result 2. *The likelihood of public good provision is significantly greater with communication than without communication (support for H2).*

Given this result and the favorable chances of high-valued public good , it is not surprising to find out that also the average group earnings are significantly higher at the 1% level in the chat treatment.[11] In each period, the groups in the communication sessions gain on average 253.58 ECU versus 179.15 ECU of the groups that are not allowed to chat. The mean difference between the two treatments is not that evident at the outset of the game, but goes beyond 110 ECU in the final periods because of the provision rate downfall in the baseline. Hence:

[11] The average group earnings in Table 4 do not consider the gain from the incentivized question.

Result 3. *The average group earnings are significantly higher in the chat sessions than in the no-communication sessions (support for H3).*

From the viewpoint of efficiency, Table 5 displays the frequency distribution of the number of contributors in the two treatments. The incidence of just one or two group members contributing (i.e., wasteful undercontribution) is significantly lower at the 1% level when communication is allowed, and amounts to only 5% of cases compared to 34% in the baseline. The striking result is that, when the opportunity to make unbinding deals is given, subjects massively prefer to coordinate on all group members contributing, these agreements being often successful despite the free-rider problem. Indeed, the incidence of all group members contributing (i.e., wasteful overcontribution) turns out to be significantly higher at the 1% level in the chat treatment (59%) than in the baseline (18%). This outcome is unprecedented in comparison with the benchmark paper (Palfrey et al., 2017). Thus:

Result 4. *The incidence of just one or two group members contributing is significantly lower in the chat treatment than in the baseline (support for H4).*

Result 5. *The incidence of all group members contributing is significantly higher in the chat treatment than in the baseline (reject H5). The four-contributor category becomes the modal class when communication is allowed.*

Finally, the incidence of groups without contributions is similar in the two treatments (two-tailed Wilcoxon rank sum test, $p = 0.69$). In the baseline, clearly, coordination on zero contributions can emerge only by chance, but in the chat treatment represents an efficient option in case the group members anticipate a low-valued public good by means of the signals. Therefore, the minimal percentages throughout the fifteen periods might be a sign of high expectations about the quality of the public

Table 4: Summary statistics at the group level

	No communication		Chat
Periods 1-5			
Provision rate	0.68		0.96***
High value		0.39	0.58
Low value		0.29	0.38
Group earnings	190.70		230.35**
No. of observations	20		20
Periods 6-10			
Provision rate	0.52		0.95***
High value		0.33	0.66
Low value		0.19	0.29
Group earnings	178.90		249.80***
No. of observations	20		20
Periods 11-15			
Provision rate	0.47		0.86***
High value		0.27	0.76
Low value		0.20	0.10
Group earnings	167.85		280.60***
No. of observations	20		20
All periods			
Provision rate	0.56		0.92***
High value		0.33	0.67
Low value		0.23	0.25
Group earnings	179.15		253.58***
No. of observations	20		20

Relative frequencies for provision rate (in turn broken down by quality of the public good); means for group earnings.
*** p-value < 0.01.
** p-value < 0.05.
* p-value < 0.10.

Table 5: Frequency distribution of number of contributors

Contributors	No communication	Chat
<u>All periods</u>		
0	0.10	0.03
1 or 2	0.34	0.05***
3	0.38	0.33
4	0.18	0.59***
Efficient outcomes (0 or 3)	0.48	0.36*
No. of observations	20	20

*** p-value < 0.01.
** p-value < 0.05.
* p-value < 0.10.

good.

4.3 Chat analysis and related discussion

Since the design is characterized by partner matching, the players are strongly encouraged to speak out and clinch long-lasting deals. We notice a not negligible tendency to devise a group strategy in the early rounds and subsequently carry it on with implicit agreement. This prevents us from interpreting the absence of relevant messages in the chat as lack of a common action profile, so that we make the realistic assumption that the twenty groups in the communication sessions reach an agreement all the time.

Table 6 suggests that the problem of overcontribution is greater than what could be observed from the mere number of group contributors: 82% of the agreements (246 out of 300 interactions) are reached on all group members contributing, even if actual overcontribution amounted only to 59% of cases. Despite the non-credible nature of these agreements and the concrete threat of free-riding, the percentage of successful deals on four contributors is surprisingly high and equal to 71% (175 out of 246). This is consistent with the fact that partner matching tends to reduce the occurrence of free-riding (Solow and Kirkwood, 2002). At the same time, the outcome might

Table 6: Agreements on the action profiles

	Agreed number of contributors			
Periods	0	1 or 2	3	4
1-5	3 (3)	0 (0)	13 (13)	69 (84)
6-10	1 (1)	0 (0)	12 (13)	63 (86)
11-15	5 (7)	0 (0)	14 (17)	43 (76)
All periods	9 (11)	0 (0)	39 (43)	175 (246)

Number of times the group members succeed in carrying out the agreed action profile in the chat treatment (number of agreements in parentheses).

be driven by the pursuit of symmetric payoffs within the group, this condition being generally conducive to coordination (Crawford et al., 2008).

Conversely, the proportion of groups with three players contributing (33%) is moderately overstated by the free-rider issue, considering that only 14% of the deals (43 out of 300) are made on three contributors. As expected, the agreements on this action profile are successfully carried out very often, up to 91% of cases (39 out of 43). Scarce convergence is detected on zero contributions, whereas none on the remaining categories.

Players perceive to be significantly more interconnected with the other group members in the chat treatment than in the baseline.[12] As possible explanation the literature puts forward the common fate hypothesis, according to which "when the members of a group are exposed to the same risk, group identification is facilitated"(Corazzini and Sugden, 2011). Further, when a player identifies with the group, theories of team reasoning claim that she will detect and implement a strategy that achieves the maximum group payoff, thereby playing her own part within such a joint action (Bacharach, 1999; Sugden, 2000, 2003). In our experimental design, partner matching allows to maximize the group earnings in every period by means of a rotation of the only one member not contributing that solves the problem of free-riding.

[12] The insight stems from the variable *Interdependency* (Wilcoxon rank sum test, $p < 0.01$), which serves as a proxy for group identity (Jackson and Smith, 1999) and is coded as in Subsection 3.2.

Therefore, in case players foresee high quality of the public good and accordingly agree on reaching the threshold, theories of team reasoning predict the implementation of such a rotation.

Notwithstanding, just 9% of the actual strategies (28 out of 300) are grounded on the rotation of the non-contributor, with a success rate equal to 96% (27 out of 28). While the scarce use of this coordination device is somewhat unexpected, high effectiveness thereof is fully predictable in that this strategy involves exclusively agreements which prevent free-riding. As an explanation, the group members might not be able to coordinate on three contributions given that it is not possible to carry out a screening of contributors. This could account for the massive overcontribution.[13] In any case, the participants in the chat are given identification numbers (i.e., "Player 1/2/3/4"), and the high success rate of rotations indicates that it is definitely feasible to take and respect turns across rounds.

In 51% of cases (154 out of 300) players devise a strategy based on the systematic sharing of all four private signals of the group, this action profile aiming to reduce uncertainty and entailing decisions period by period. In 97% of these interactions (150 out of 154) all the group members honestly report the true values and 83% of the signal-based strategies (128 out of 154) are successfully implemented. Hence, free-riders do not overstate the public good value, despite their clear incentive pointed out in the literature (Crawford and Sobel, 1982).

The remaining 39% of the observations (118 out of 300) are mainly represented by interactions where only some of the group members disclose private information. The related lower success rate (58%) might be justified by lack of a well-defined strategy. Accordingly:

Result 6. *In the chat treatment, rotation-based strategies are less popular than signal-based strategies (reject H6), which constitute the modal class.*

[13] We thank an anonymous referee for proposing this argument.

All in all, the groups do not maximize earnings but prefer to pursue symmetric payoffs and minimize ambiguity through the sharing of signals period by period, this strategy being inefficient because not feasible jointly with the rotation of the non-contributor.

To corroborate this evidence and gain new insights, as a final step we perform a content analysis of the text chat and code the individual messages into six mutually-exclusive categories, in addition to a residual class.[14]

Table 7: Content analysis: messages by category

Category	Messages
(1) Own signal/others' signals/signal-based strategy suggestion	26.15
(2) Rotation-based strategy suggestion	2.06
(3) Other strategy suggestion	8.55
(4) Confirmation	14.32
(5) Informative statements/explaining something to group members	13.43
(6) Irrelevant/junk	20.77
Others (group identity/lying/trust/past coordination outcomes)	14.72

Percentage of messages falling within category.

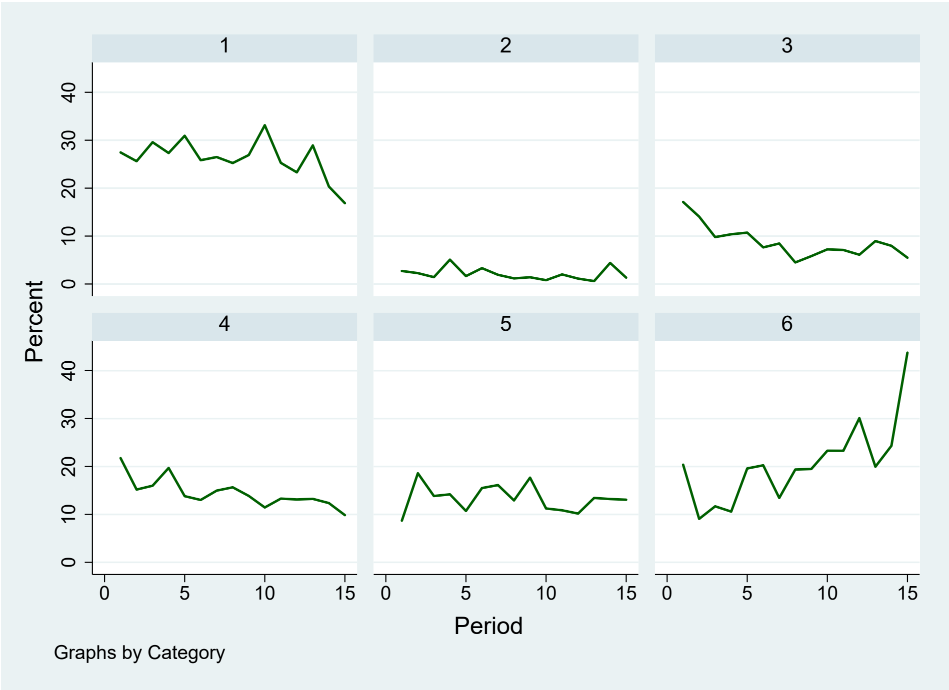
Table 7 highlights once again that the uncertainty-related discussion outweighs any other topic: category (1) contains 26% of the chat messages, while rotation-based (2%) and alternative strategies (9%) are infrequently debated.[15]. The numbers of confirmation and informative statements turn out to be in line with those in Palfrey et al. (2017), implying that subjects in general understand the game. The relatively high percentage of irrelevant messages (21%) is arguably inherent in the dynamics

[14]Following Palfrey et al. (2017), in Appendix B we report a table displaying examples of sentences used in the experiment for each category.

[15]Nonetheless, it is worthwhile to draw attention to the unique group carrying on the rotation throughout the fifteen rounds, where category (2) covers 21% of the group messages. This group outperforms all the others when it comes to making profits, as well as is able to respect turns all the time thanks to participant identification numbers. Notably, according to content analysis the four group members make no use of private signals, category (1) reaching barely 2% of the group messages. This occurrence stresses the need to shed further light on the role of uncertainty in the adoption of rotation schemes.

of partner matching that can be better appreciated in Figure 3. Indeed, examining how communication changes over time, we observe a progressive increment in the messages of category (6), which gradually crowd out strategy suggestion and resulting confirmation statements. Basically, small talk proliferates as group members get acquainted with each other. The remaining chat messages just bring up a wide array of side topics (e.g., group identity, lying, trust) that are included in an unnumbered residual category.[16]

Figure 3: Content analysis: messages by period



5 Conclusions

In this paper we offer experimental evidence on the impact of communication on the provision of public goods whose value is ambiguous. We draw inspiration from school

[16]The time pattern of the residual category is reported in Appendix B.

construction in a developing country by a NGO as a case where the quality of the final output cannot be inferred in advance. The current work purports to analyze the dynamics of cooperation under an innovative concept of uncertainty, thereby testing the most recent conclusions related to low-quality public goods (Böhm et al., 2016), efficiency of communication (Palfrey et al., 2017) and team reasoning as main theory for the explanation of coordination game data (Bardsley and Ule, 2017).

The analysis through dynamic panel data models emphasizes that unexpected low-valued public goods can actually undermine future willingness to contribute only when communication is allowed. This outcome is in line with Böhm et al. (2016), who observe a significant behavioral change after vaccine-adverse events in an Interactive Vaccination Game. Given the scarce use of dynamic models in the experimental literature, we also reaffirm the effectiveness of this technique in bringing new results to light and support the spread thereof for the study of repeated games. In any case, the benefits of communication prevail over the drawbacks, since coordination leads to higher public good provision and reduces wasteful undercontribution in the chat treatment. This result appears to be complementary to Cox and Stoddard (2018), who in a similar context of common-value public goods investigate truthful cheap talk and find only limited detrimental effects of communication due to incentives to lie.

In addition, we detect unprecedented inefficiency in the form of massive overcontribution, given that subjects tend to neglect the free-rider problem. Chat analysis reveals that the group agreements on all members contributing are often successful, arguably thanks to pronounced group identity as well as players' willingness to pursue symmetric payoffs. Future research might test the unusual efficacy of such deals under lower public goods thresholds, hopefully drawing novel insights to discourage free-riders.

Despite the presence of a valid coordination device as the rotation of the non-contributor, the group members prefer the implementation of strategies based on the systematic disclosure of all four private signals, or in other words, the minimization of uncertainty is favored over the maximization of the group earnings. Considering the trade-off between the two goals, this manuscript also spotlights cases where theories of team reasoning have limited explanatory power. As a result, we encourage further variations of this TPG game to discover whether uncertainty makes satisficing more salient than optimizing, for instance by replacing ambiguity about the public good value with the stated probabilities.

To conclude, regarding possible applications we stress the fact that the experiment is not framed, so that a number of compatible environments can benefit from the results (e.g., prevention of epidemics and natural disasters). On the other hand, it is harder to say how serious the inefficiency from overcontribution is, the answer arguably depending on the various real-world situations. Definitely, further research is needed to shed light on this aspect.

Acknowledgements

We gratefully acknowledge Fabio Tufano, Maria Bigoni, Nikolaos Georgantzis, Antonio Filippin, Annarita Colasante and three anonymous reviewers for helpful comments on early versions of this manuscript, as well as the *Spanish Ministerio de Ciencia, Innovación y Universidades* (grant RTI2018-096927-B-I00) and the *Universitat Jaume I* (grant UJI-B2018-76) for financial support. We also thank the participants in the PhD workshops held at Università degli Studi dell'Insubria in Varese (May 3, 2018) and Universitat Jaume I in Castellón (May 14, 2018), as well as in the 11^{th} Maastricht Behavioral Experimental Economics Symposium in Maastricht (June 4, 2018), the 15^{th} Workshop on Social Economy for Young Economists in Forlí (June 8-9, 2018), the 9^{th} International Conference of the French Association of Experimental Economics in Nice (June 14-15, 2018) and the Vechta Workshop on Social Behavior in Vechta (January 24-25, 2019) for useful remarks. This paper supersedes our earlier manuscript "Communication in a threshold public goods game with ambiguity: Anomalies and regularities".

References

Agastya, M., Menezes, F., and Sengupta, K. (2007). Cheap talk, efficiency and egalitarian cost sharing in joint projects. *Games and Economic Behavior*, 60(1):1–19.

Alberti, F. and Cartwright, E. J. (2016). Full agreement and the provision of threshold public goods. *Public Choice*, 166(1):205–233.

Artinger, F., Fleischhut, N., Levati, M. V., and Stevens, J. R. (2012). Cooperation in a risky environment: Decisions from experience in a stochastic social dilemma. *Jena Economic Research Papers, 2012/047.*

Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, 53(2):117–147.

Bailey, R. W., Eichberger, J., and Kelsey, D. (2005). Ambiguity and public good provision in large societies. *Journal of Public Economic Theory*, 7(5):741–759.

Balafoutas, L., Kocher, M. G., Putterman, L., and Sutter, M. (2013). Equality, equity and incentives: An experiment. *European Economic Review*, 60:32–51.

Bardsley, N. and Ule, A. (2017). Focal points revisited: Team reasoning, the principle of insufficient reason and cognitive hierarchy theory. *Journal of Economic Behavior & Organization*, 133:74–86.

Baron-Cohen, S. and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2):163–175.

Barrett, S. and Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences*, 109(43):17372–17376.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33(1):1–27.

Bicchieri, C. and Lev-On, A. (2007). Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. *Politics, Philosophy & Economics*, 6(2):139–168.

Blanco, M., Engelmann, D., Koch, A. K., and Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4):412–438.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Bochet, O., Page, T., and Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60(1):11–26.

Böhm, R., Betsch, C., and Korn, L. (2016). Selfish-rational non-vaccination: Experimental evidence from an interactive vaccination game. *Journal of Economic Behavior & Organization*, 131:183–195.

Bortolotti, S., Casari, M., and Pancotto, F. (2015). Norms of punishment: Experiments with students and the general population. *Economic Inquiry*, 53(2):1207–1223.

Brañas-Garza, P., Bucheli, M., Espinosa, M. P., and García-Muñoz, T. (2013). Moral cleansing and moral licenses: Experimental evidence. *Economics & Philosophy*, 29(2):199–212.

Brañas-Garza, P., Bucheli, M., García-Muñoz, T., et al. (2011). Dynamic panel data: A useful technique in experiments. *Technical Report 10/22, Department of Economic Theory and Economic History of the University of Granada.*

Brosig, J. (2002). Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior & Organization*, 47(3):275–290.

Calford, E. (2017). Uncertainty aversion in game theory: Experimental evidence. *Purdue University, Department of Economics, Working Paper No. 1291*.

Caporael, L. R., Dawes, R. M., Orbell, J. M., and Van De Kragt, A. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, 12(4):683–699.

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

Cherry, T., Howe, E. L., and Murphy, J. (2015). Sharing as risk pooling in a social dilemma experiment. *Ecology and Society*, 20(1).

Cooper, R., DeJong, D. V., Forsythe, R., and Ross, T. W. (1992). Communication in coordination games. *The Quarterly Journal of Economics*, 107(2):739–771.

Corazzini, L. and Sugden, R. (2011). Common fate, game harmony and contributions to public goods: experimental evidence. *International Review of Economics*, 58(1):43–52.

Cox, C. A. (2015). Cursed beliefs with common-value public goods. *Journal of Public Economics*, 121:52–65.

Cox, C. A. and Stoddard, B. (2018). Common-value public goods and informational social dilemmas. *Working Papers 18-04, Department of Economics, Appalachian State University*.

Crawford, V. P., Gneezy, U., and Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98(4):1443–1458.

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

Dawes, R. M., Orbell, J. M., Simmons, R. T., and Van De Kragt, A. (1986). Organizing groups for collective action. *American Political Science Review*, 80(4):1171–1185.

Dhami, S. (2016). *The foundations of behavioral economic analysis*. Oxford University Press.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Eichberger, J. and Kelsey, D. (2002). Strategic complements, substitutes, and ambiguity: the implications for public goods. *Journal of Economic Theory*, 106(2):436–466.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

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

Gangadharan, L. and Nemes, V. (2009). Experimental analysis of risk and uncertainty in provisioning private and public goods. *Economic Inquiry*, 47(1):146–164.

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

Güth, W., Levati, M. V., and Soraperra, I. (2015). Common and private signals in public goods games with a point of no return. *Resource and Energy Economics*, 41:164–184.

Harrison, G. W., Martínez-Correa, J., and Swarthout, J. T. (2014). Eliciting subjective probabilities with binary lotteries. *Journal of Economic Behavior & Organization*, 101:128–140.

Jackson, J. W. and Smith, E. R. (1999). Conceptualizing social identity: A new framework and evidence for the impact of different dimensions. *Personality and Social Psychology Bulletin*, 25(1):120–135.

Keenan, D. C., Kim, I., and Warren, R. S. (2006). The private provision of public goods under uncertainty: A symmetric-equilibrium approach. *Journal of Public Economic Theory*, 8(5):863–873.

Kelsey, D. and Le Roux, S. (2017). Dragon slaying with ambiguity: theory and experiments. *Journal of Public Economic Theory*, 19(1):178–197.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., and David, A. S. (2004). Measuring empathy: reliability and validity of the empathy quotient. *Psychological Medicine*, 34(5):911–920.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. and Roth, A., editors, *Handbook of Experimental Economics*. Princeton: Princeton University Press.

Liebrand, W. B. (1984). The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixed-motive game. *European Journal of Social Psychology*, 14(3):239–264.

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

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *Economic Journal*, 106:817–845.

Ostrom, E., Gardner, R., and Walker, J. (1994). *Rules, games, & common-pool resources*. University of Michigan Press.

Palfrey, T. and Rosenthal, H. (1991). Testing for effects of cheap talk in a public goods game with private information. *Games and Economic Behavior*, 3(2):183–220.

Palfrey, T., Rosenthal, H., and Roy, N. (2017). How cheap talk enhances efficiency in threshold public goods games. *Games and Economic Behavior*, 101:234–259.

Parks, C. D. (1994). The predictive ability of social values in resource dilemmas and public goods games. *Personality and Social Psychology Bulletin*, 20(4):431–438.

Solow, J. L. and Kirkwood, N. (2002). Group identity and gender in public goods experiments. *Journal of Economic Behavior & Organization*, 48(4):403–412.

Sonnemans, J., Schram, A., and Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, 34(1):143–161.

Sugden, R. (2000). Team preferences. *Economics & Philosophy*, 16(2):175–204.

Sugden, R. (2003). The logic of team reasoning. *Philosophical Explorations*, 6(3):165–181.

Tavoni, A., Dannenberg, A., Kallis, G., and Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, 108(29):11825–11829.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126(1):25–51.