Journal of the Economic Science Association (2023) 9:308–318
https://doi.org/10.1007/s40881-023-00139-1

METHODOLOGY PAPER

Cooperation and confusion in public goods games: confusion cannot explain contribution patterns

Armin Granulo[1] · **Rudolf Kerschreiter**[2] · **Martin G. Kocher**[3,4]

Received: 14 November 2021 / Revised: 13 July 2023 / Accepted: 14 July 2023 /
Published online: 10 August 2023

Abstract

People behave much more cooperatively than predicted by the self-interest hypothesis in social dilemmas such as public goods games. Some studies have suggested that many decision makers cooperate not because of genuine cooperative preferences but because they are confused about the incentive structure of the game—and therefore might not be aware of the dominant strategy. In this research, we experimentally manipulate whether decision makers receive explicit information about which strategies maximize individual income and group income or not. Our data reveal no statistically significant effects of the treatment variation, neither on elicited contribution preferences nor on unconditional contributions and beliefs in a repeated linear public goods game. We conclude that it is unlikely that confusion about optimal strategies explains the widely observed cooperation patterns in social dilemmas such as public goods games.

Keywords Public goods game · Cooperation · Experiment · Confusion

JEL Classification C91 · H41

✉ Martin G. Kocher
martin.kocher@univie.ac.at

Armin Granulo
armin.granulo@tum.de

Rudolf Kerschreiter
rudolf.kerschreiter@fu-berlin.de

[1] TUM School of Management, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany

[2] Division of Social, Organizational, and Economic Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

[3] Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria

[4] University of Gothenburg, Gothenburg, Sweden

1 Introduction

A vast number of laboratory and field studies have shown that many people contribute voluntarily to the provision of public goods, even when it is not in their monetary interest (e.g., Chaudhuri, 2011; Fehr & Fischbacher, 2003; Zelmer, 2003). The declining pattern of contributions over time in these social dilemmas is consistent with two behavioral explanations. First, conditional cooperation (or reciprocity) has been identified as being important for voluntary contributions to public goods; i.e., many decision makers contribute to public goods when others also contribute or are expected to do so (Fischbacher et al., 2001; Fischbacher & Gächter, 2010; Thöni & Volk, 2018). Second, confusion (or decision errors) have been invoked as an alternative explanation for the prevalence of voluntary contributions; i.e., many decision makers, particularly in laboratory experiments, are supposed to contribute to a public good not out of a preference motive or out of a reciprocity norm, but instead because they misunderstand the incentives of the game and therefore are unaware how to correctly pursue their self-interest (e.g., Andreoni, 1995; Burton-Chellew & West, 2013; Ferraro & Vossler, 2010; Houser & Kurzban, 2002). Confusion might be especially relevant in one-shot interactions or at the start of repeated interactions, and the observed decay in contribution levels, if it is due to learning, is consistent with this explanation. For example, Burton-Chellew et al. (2016), using the strategy method elicitation developed by Fischbacher et al. (2001), report that decision makers exhibit the same conditional contribution pattern, irrespectively of whether they interact with humans or computers, which seems to corroborate the second explanation based on confusion.

In this paper, we report on a novel way of testing whether confusion about optimal strategies, i.e., optimally implementing one's preference, could be an important driver of voluntary contributions in a laboratory public goods game. Given the importance of public goods games in analyzing social dilemmas and developing policy-relevant designs and incentives for problems outside the laboratory (see, for instance, Schmidt & Ockenfels, 2021, for an application in climate policy), it seems relevant to assess the internal validity of the main paradigm used in experimental research. To this end, we experimentally vary in a linear public goods game whether decision makers receive information about the individually optimal strategy and the socially optimal strategy or not, and we analyze how this information affects elicited contribution preferences and cooperation in a public goods game. We believe that an assessment of whether experimental participants fully understand their strategic options and the relevant incentives provides us with the most direct test of the confusion hypothesis.

2 Experimental design and procedures

Our experimental design builds on the standard voluntary contribution mechanism with the following linear payoff function:

$$\pi_i = 20 - g_i + 0.5 \sum_{j=1}^{3} g_i, \tag{1}$$

where g_i denotes the contribution of participant i to the public good. Each group consists of $n=3$ randomly assigned participants, and each participant receives an endowment of 20 points. The marginal per capita return (MPCR) from investing in the public good is 0.5 and the social return is 1.5. The parameters are all known by participants. Assuming that participants are rational and selfish payoff maximizers, these parameters guarantee that it is individually optimal to contribute zero. From a social or efficiency perspective, they guarantee that it is collectively optimal to contribute the entire endowment. Hence, the setup and the parameters imply a social dilemma.

Participants were randomly assigned to one of four treatments in a 2 (information: standard vs. optimal strategies)×2 (control questions: incentivized vs. not incentivized) between-subjects design (see Table 1). In the *standard information treatment*, participants received the standard instructions that explained the public goods problem (see Online Appendix A; online link is in the acknowledgments). In the *optimal strategies information treatment*, participants received the same instructions, but with an additional paragraph that was explicit about which strategies maximize their individual income and their group's income. Specifically, participants received the information that their individual income is maximized by contributing zero to the public good, regardless of the behavior of the other group members, and why this is the case (see Online Appendix A); i.e., we explained strategic dominance. Participants additionally received the information that their group's income is maximized by contributing one's entire endowment to the public good if everybody does so, and why this is the case (see Online Appendix A). Participants were also explicitly informed that in case they contribute more to the public good than their group members, the other group members benefit more from their contributions and end up earning more, i.e., we explained the sucker's payoff.

At the beginning, participants learned that the experiment consists of four parts.[1]

Part 1 Participants were asked to answer 16 standard control questions in four separate blocks (see Online Appendix B). We often use these or similar questions in related public goods experiments to ensure a basic understanding. In the *control questions incentivized treatment*, participants could earn a bonus. They were told that, after completing all questions, one question would be randomly chosen and, if answered correctly, would result in a bonus of 12 experimental points. In the *control questions not incentivized treatment*, participants could not earn a bonus. After completing all questions, all participants received information for every question about the correct answer, their actual answer, and if their answer was correct or wrong.

Part 2 We then elicited contribution preferences using the strategy method of Fischbacher et al. (2001), validated for repeated interactions in Fischbacher and Gächter

[1] Instructions can be found in the Online Appendix. Part 4 elicited fairness norms. For reasons of succinctness, we do not report the results here.

Table 1 Summary statistics for the four treatments

Treatment	Part 1: control questions	Part 2: strategy method		Part 3: repeated public goods game		*N*
	Av. corr. contr. questions	Av. uncond. contrib.	Av. slopes	Av. contrib.	Av. beliefs	
Control questions incentivized and optimal strategies information	0.80 (0.30)	7.79 (6.15)	0.60 (0.58)	7.28 (5.50)	7.73 (4.85)	24
Control questions not incentivized and optimal strategies information	0.82 (0.21)	8.29 (6.89)	0.49 (0.46)	7.70 (7.01)	9.00 (6.47)	21
Control questions incentivized and standard information	0.73 (0.31)	10.08 (7.04)	0.47 (0.54)	10.08 (4.91)	11.08 (3.27)	24
Control questions not incentivized and standard information	0.76 (0.28)	8.38 (6.32)	0.54 (0.51)	8.19 (6.03)	10.28 (5.16)	24
Treatments with optimal strategies information (pooled)	0.81 (0.26)	8.02 (6.44)	0.55 (0.52)	7.47 (6.18)	8.32 (5.63)	45
Treatments with standard information (pooled)	0.75 (0.30)	9.23 (6.68)	0.50 (0.52)	9.14 (5.53)	10.68 (4.28)	48

Av. corr. contr. questions = average correct control questions; Av. uncond. contrib. = average unconditional contributions; Av. contr. = average contribution; *N* = number of participants. Standard deviations in parentheses

Slopes are calculated as the slope coefficient of an individual regression with own contribution as the dependent variable and average others' contributions as the independent variable

(2010). Group members first make an unconditional contribution to the public good, which is a single integer number that satisfies $0 \leq g_i \leq 20$. Thereafter, group members make a conditional contribution for each of the 21 possible rounded averages from 0 to 20 (i.e., they submit a contribution schedule). Both the unconditional as well as the conditional contribution are potentially payoff relevant (for the way of how to incentivize both, see Fischbacher et al., 2001). Participants did not receive any information about other participants' decisions at the end of Part 2.

Part 3 Finally, participants played ten periods of a repeated public goods game with partner matching (i.e., with constant group composition). We emphasized that the group composition was determined randomly and thus most likely different from the previous strategy method decisions in Part 2. In each period, we elicited beliefs about the other group members' average contributions. At the end of each period, group members were informed about their group members` average contributions and their own payoffs. To avoid hedging, at the end of the session two periods were randomly selected for payment: In one period, the outcome of the public goods game was payoff relevant, and in another one, beliefs were incentivized.

We ran the experiment in the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA), using the software z-Tree (Fischbacher, 2007) and the organizational software ORSEE (Greiner, 2015). In total, 93 undergraduates took part in the experiment with average earnings of €19 (including a show-up fee of €4).

3 Experimental results

Table 1 provides summary statistics for the three parts of the four treatments and shows the number of independent observations per treatment. Our analysis reveals that incentivizing control questions does not significantly affect participants' behavior in any of the three parts (see Online Appendix C for a detailed discussion of the results). Since there is no effect of the incentivizing control questions, we pool the data and compare results for the two information treatments in the remainder of the paper.

In this context, it is also important to see whether different experimental instructions influence the number of correct answers. As a matter of fact, information about optimal strategies has neither a significant effect on how many correct answers participants give [$p = 0.313$; two-sided Mann–Whitney-U (MWU) test; all p values $\geq$ 0.084 (MWU) on the level of the 16 individual control questions], nor on how many participants answer all control questions correctly ($p = 0.804$; two-sided Chi-square test). Hence, information about optimal strategies does not affect how well participants do in answering control questions.

3.1 Treatment information: standard instructions vs. optimal strategies instructions

We observe that information about the optimal strategies does not affect the distribution of contribution preferences elicited by the strategy method ($p = 0.366$;

Table 2 Distribution of player types for the two information treatments

Treatment	Free rider	Conditional cooperators	Hump-shaped/others
Optimal strategies information	17.8% ($N=8$)	66.7% ($N=30$)	15.5% ($N=7$)
Standard information	12.5% ($N=6$)	60.4% ($N=29$)	27.1% ($N=13$)
Chi^2 test	$p=0.366$		

There are $N=2$ humped-shaped contributors in the optimal strategy information treatment and $N=4$ in the standard information treatment. To avoid that a cell has fewer than five entries, we pool hump-shaped and others types (see Kocher et al., 2015)

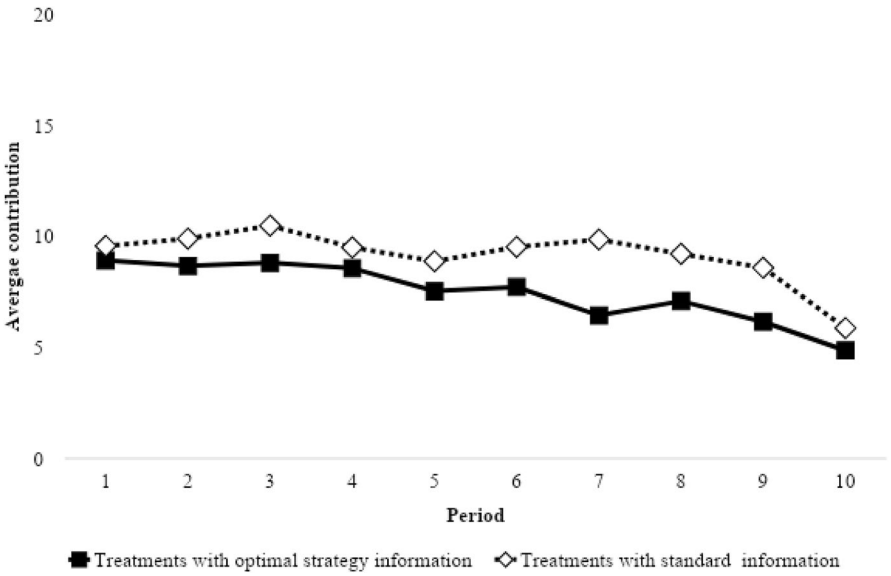

Fig. 1 Average contributions over ten periods in Part 3

two-sided Chi-square test; definitions of types according to Fischbacher et al., 2001; see Table 2). Especially, the relative frequency of conditional cooperators is almost identical. Further, Mann–Whitney-U tests do not reveal any significant differences in slopes of the conditional cooperation schedules[2] ($p=0.561$) or mean unconditional contributions in Part 2 ($p=0.363$) for the two treatments. Thus, we conclude that decision makers exhibit the same elicited preferences for cooperation, regardless of whether they receive standard instructions or instructions that explain the payoff consequences of all strategies in detail. Confusion does not seem to play a major role.

In Fig. 1, we show the dynamics of contributions over the ten periods of Part 3 for the two information treatments. We find that information about optimal strategies

[2] The slope coefficient is calculated as the slope coefficient of an individual regression with own contribution as the dependent variable and average others' contributions as the independent variable.

Table 3 Regression models of contributions in the repeated public goods game in period 1

Model	Dependent variable: contribution			
	1	2	3	4
Optimal strategies information	–0.651 (1.460)	0.596 (0.800)	0.580 (0.801)	–0.003 (1.036)
Belief		1.011*** (0.036)	0.930*** (0.080)	0.980*** (0.096)
Predicted contribution			0.100 (0.080)	–0.004 (0.088)
Optimal strategies information × belief				–0.069 (0.154)
Optimal strategies information × predicted contribution				0.183 (0.154)
Constant	9.563*** (0.925)	–0.354 (0.658)	–0.272 (0.686)	–0.024 (0.847)
Observations	93	93	93	93
R-squared	0.002	0.719	0.723	0.728

Coefficients from OLS regressions with robust standard errors reported in parentheses

***p < 0.01, **p < 0.05, *p < 0.1

does not affect mean contributions ($p = 0.304$; MWU; group means as independent observations) and mean beliefs ($p = 0.149$; MWU; group means as independent observations) in the repeated public goods game (Part 3). To investigate this relationship, we run a regression of participants' contributions in the public goods game in period 1, using OLS regressions (see Models 1–4; Table 3) and averaged over periods 1–10 using GLS regressions (see Models 5–8; Table 4). The models include different predictors. Models 1 and 5 include an information treatment dummy; Models 2 and 6 additionally include beliefs on others' contributions; Models 3 and 7 add predicted contributions (i.e., contributions based on elicited preferences from Part 2 and the belief); and Models 4 and 8 further add interaction terms of the information treatment dummy with beliefs and predicted contributions as well as period (only Model 8). The analysis reveals that information about the optimal strategies has no significant main effect on average contributions, neither in period 1 (see Models 1–4) nor overall (see Models 5–8). Moreover, information about optimal strategies does not significantly interact with beliefs and predicted contributions, neither in period 1 (see Models 1 and 4) nor overall (see Models 5 and 8), nor with period (see Model 8). Based on these results, we conclude that participants' contributions in a repeated public goods game, and their beliefs about others' contributions, do not change when they receive extended information about which strategies maximize individual and group payoffs.

3.2 Statistical power

Participants clearly provide positive contributions to the public good, even with extended information. Hence, our main empirical conclusion is safely established. Since we get a null result with regard to our treatment variation, however, it is

Table 4 Regression models of contributions in the repeated public goods game in periods 1–10

Model	Dependent variable: contribution			
	5	6	7	8
Optimal strategies information	–1.662 (1.976)	–0.129 (0.979)	–0.051 (0.884)	–1.511 (1.158)
Belief		0.651*** (0.068)	0.521*** (0.094)	0.429*** (0.128)
Predicted contribution			0.236*** (0.089)	0.290** (0.143)
Optimal strategies information × belief				0.182 (0.170)
Optimal strategies information × predicted contribution				–0.103 (0.168)
Optimal strategies information × period				0.073 (0.119)
Period	–0.349*** (0.096)	–0.202*** (0.057)	–0.201*** (0.060)	–0.229*** (0.075)
Constant	11.052*** (1.147)	3.293*** (0.818)	2.956*** (0.786)	3.689*** (0.854)
Observations	930	930	930	930
Number of individuals	93	93	93	93

Coefficients from GLS regressions with standard errors in parentheses, random effects on the participant level and standard errors clustered at the matching group level

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

relevant to give an impression of the statistical power that we operate with. Given our sample size reported in Table 1 and our data, we can perform ex post power calculations for null results, as recommended by Nikiforakis and Slonim (2015). Specifically, we can calculate the minimum treatment effect that we could have detected with 80% power at a 5% significance level.[3] For unconditional contributions elicited in the strategy method (Part 2), this analysis revealed that the minimum detectable effect of the information treatment that we could have detected is 3.81 (i.e., a decrease from an average of 9.23 in the standard information treatment to an average of 5.42 in the information about optimal strategies treatment). Taking into account covariates (i.e., slopes and intercepts of the reciprocity functions), the minimum detectable effect size reduces to 3.25. For contributions in a repeated linear public goods game (Part 3), the minimum detectable effect size of the information treatment was 4.07 in period 1 and 3.37 for periods 1–10. Taking into account covariates (i.e., a treatment dummy for incentivized control questions, beliefs, and predicted contributions), this minimum detectable effect size reduces to 2.18 in period 1 and to 1.60 for periods 1–10. In our experiment, the observed effect size of the information treatment on unconditional contributions is generally below these thresholds [observed effects sizes: 1.21 (Part 2), 0.65 in period 1 and 1.66 over periods 1–10 (Part 3)]. When interpreting the significance of our treatment differences, these aspects should be taken into account.

4 Conclusion

Players behave much more cooperatively than predicted by the self-interest hypothesis in social dilemmas such as public goods games. Some studies have suggested that the decay pattern in repeated public goods games, and the results from experiments in which humans interact with computers, are indicative of the hypothesis that many decision makers cooperate not because of following genuine cooperative preferences, but because of confusion about the incentive structure of the game, and thus not being aware of the dominant strategy in the game. The decay is interpreted by these studies as an indication of learning.

In this paper, we experimentally manipulate in a linear public goods game whether decision makers receive explicit information about the individually optimal strategy and the socially optimal strategy or not, and we analyze how this information affects elicited contribution preferences and cooperation. More precisely, we discuss the payoff consequences of strategies in the game explicitly and in detail in the experimental instructions in one treatment and use standard instructions in another. While the individually optimal strategy that we used in the instruction holds only under common knowledge in Part 3 of the experiment, we think that such common knowledge might have been implemented through reading the instructions out loud. However, in Part 2, the individually optimal strategy described in the instruction holds strictly in any case.

[3] See Bloom (1995) and Bloom et al. (2007) for power calculations.

Our data do not reveal statistically significant effects of the treatment variation on participants' understanding of the task (Part 1), elicited contribution preferences (Part 2), or unconditional contributions and beliefs in a repeated linear public goods game (Part 3). Contributions are positive in any case and the size of the contributions is similar to related experiments. We conclude that it is unlikely that confusion about optimal strategies is a relevant explanation for the widely observed cooperation patterns in social dilemmas such as public goods games.

Why do other approaches obtain results in favor of the confusion hypothesis? A set of studies on human–computer interactions in public goods games or prisoner's dilemmas (e.g., Burton-Chellew et al., 2016) finds only small differences between choices against another human player (or: other human players) and a computer algorithm. We think that it would be interesting and worthwhile to conduct an experiment that systematically varies the information that participants receive about the algorithm used by the computer player (or: by the computer players). Such an experiment could rigorously establish whether different levels of information matter in human–computer interaction. Our results would indicate that for human–computer interactions, making the optimal strategy against the algorithm clearer, should result into play closer to the dominant strategy. However, only a rigorous experiment can establish such a claim empirically.

Acknowledgements Financial support from the Ideenfonds of the University of Munich (financed through the excellence initiative) and the Economics Department of the University of Munich is gratefully acknowledged. We thank Maria Bigoni as the editor in charge and two anonymous referees for very helpful suggestions that improved the paper a lot. The replication material and other supplementary material for the study is available at https://doi.org/10.17605/OSF.IO/6DFZ4.

Funding Open access funding provided by University of Vienna.

Data availability Data are available in the replication package—link is in the acknowledgements.

References

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *The American Economic Review, 85*, 891–904.

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review, 19*(5), 547–556.

Bloom, H. S., Richburg-Hayes, L., & Black, A. R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis, 29*(1), 30–59.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences, 113*(5), 1291–1296.
Burton-Chellew, M. N., & West, S. A. (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences, 110*(1), 216–221.
Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics, 14*(1), 47–83.
Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.
Ferraro, P. J., & Vossler, C. A. (2010). The source and significance of confusion in public goods experiments. *The BE Journal of Economic Analysis & Policy, 10*(1), 53.
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178.
Fischbacher, U., & Gachter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review, 100*(1), 541–556.
Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*(3), 397–404.
Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114–125.
Houser, D., & Kurzban, R. (2002). Revisiting kindness and confusion in public goods experiments. *American Economic Review, 92*(4), 1062–1069.
Kocher, M. G., Martinsson, P., Matzat, D., & Wollbrant, C. (2015). The role of beliefs, trust, and risk in contributions to a public good. *Journal of Economic Psychology, 51*, 236–244.
Nikiforakis, N., & Slonim, R. (2015). Editors' preface: Statistics, replications and null results. *Journal of the Economic Science Association, 1*(2), 127–131.
Schmidt, K. M., & Ockenfels, A. (2021). Focusing climate negotiations on a uniform common commitment can promote cooperation. *Proceedings of the National Academy of Sciences, 118*(11), e2013070118.
Thöni, C., & Volk, S. (2018). Conditional cooperation: Review and refinement. *Economics Letters, 171*, 37–40.
Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics, 6*(3), 299–310.