Using the Strategy Method and Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments

Simon Gächter[a] Diego Marino Fages[b]

January 10, 2023

Abstract

In this paper we disentangle the role of cooperative preferences and beliefs for explaining MPCR and group size effects in public goods games. To achieve this, we use the ABC approach, which explains cooperation as a function of cooperative attitudes and beliefs. We measure cooperative attitudes using the incentive-compatible strategy method as introduced by Fischbacher et al. (2001, FGF) for small groups. However, to keep the strategy method incentive-compatible across all group sizes, we need a strategy method that is scalable to any group size. Our scalable strategy method, which selects conditional and unconditional contributions with equal probability, performs well in in-sample and out-of-sample predictions of cooperation, and also compared to FGF. We find that preference types are similar across group sizes of 3 and 9 and MPCRs of 0.4 and 0.8. Further experiments comparing group sizes of 3 and 30 again find similar distributions of conditional preferences. By using the ABC approach, we show that controlling for elicited preferences, differences in cooperation levels observed across the various games are mostly due to differences in beliefs.

JEL: C92, H41

Keywords: Public Goods, Group Size, MPCR, Strategy Method, ABC approach, Experiments

[a]University of Nottingham, IZA and CESifo. E-mail: simon.gaechter@nottingham.ac.uk.

[b]Corresponding author. University of Nottingham. E-mail: diego.marinofages@nottingham.ac.uk, diegomarinofages@gmail.com.

1 Introduction

Public good games (PGGs) are one of the classic games to study cooperation in social dilemmas. PGGs have been used to investigate many aspects of voluntary cooperation (for surveys see Chaudhuri (2011); Fehr and Schurtenberger (2018); Gächter and Herrmann (2009); Ledyard (1995); Zelmer (2003); Drouvelis (2021)). In this paper, we revisit a classic question in PGG research by separately measuring preferences and beliefs: the role of group size and the marginal per capita return (MPCR) for cooperation.

One important insight that will be crucial for our analysis of MPCR and group size effects is that contributions are correlated with beliefs about others' contributions (e.g., Croson (2007); Fischbacher and Gächter (2010)). To go beyond mere correlations, Fischbacher et al. (2001) (henceforth FGF) introduced a tool based on the strategy method that measures the causal influence of beliefs about others' contributions on own contribution. The FGF method measures people's preferences for cooperation conditional on how much others contribute to the public good and allows classifying people into preference types as conditional cooperators, free riders and others.

FGF find that most people can be characterized as conditional cooperators or free rider types. The distribution of types (in particular conditional cooperators) has been shown to be similar across comparable studies (Thöni and Volk, 2018); mostly robust to the (mis-)understanding of incentives (Gächter et al., 2022; Fosgaard et al., 2017); stable over time (Gächter et al., 2022; Volk et al., 2012); mostly similar across different cooperation games (Mullett et al., 2020; Eichenseer and Moser, 2020) and similar across different cultures (Weber et al., 2021). Moreover, the ABC approach - using people's conditional preferences (also called *a*ttitudes, dispositions) together with their *b*eliefs - can explain individual *c*ontributions to public goods fairly well (Fischbacher and Gächter, 2010; Fischbacher et al., 2012; Gächter et al., 2017, 2022; Isler et al., 2021), which makes the ABC approach a promising methodology for analyzing MPCR and group size effects in cooperation.

A limitation of the existing evidence on the distribution of conditionally cooperative attitudes (preference types) as well as on the predictive power of the ABC approach to explain contributions is, however, that it mostly comes from very similar group sizes of 3 and 4 group members, with one exception (Mill and Theelen, 2019) who also studied group sizes of 5 and 7.[1] As far as we are aware, there is no evidence on the distribution of cooperative types across substantially different group sizes. Similarly, not much is known about how the MPCR affects conditionally cooperative preferences because most

[1]Relatedly, in the Cooperation Databank (a recent effort by Spadaro et al. (2022) to collect papers on cooperation in order to facilitate meta-analyses) only 8 out of 137 published papers related to continuous PGGs (with no deception) mention group size in the treatment.

studies use MPCRs of 0.4 or 0.5. Finally, to our knowledge, there is no evidence on how well the ABC approach predicts cooperation in larger groups. Hence, two main goals of our paper are, first, to measure how group size and MPCR affect attitudes, beliefs, and contributions; and, second, to apply the ABC approach for analyzing MPCR and group size effects in cooperation.

We present two studies based on one-shot games. In Study 1 (n=936), we use the strategy method and the ABC approach to study group sizes of 3 and 9 and MPCRs of 0.4 and 0.8, in a factorial design. In Study 2 (n=600) we extend group size to 30 members. We use the ABC approach to be able to disentangle preference and belief effects: if we observe different contributions across different group sizes or different MPCRs, it could be because (i) cooperative attitudes vary with group size (or MPCR) or (ii) because individuals respond to beliefs that change with group size (or MPCR) or (iii) both cooperative attitudes and beliefs are affected by group size (or MPCR).

Using the strategy method as developed by FGF to measure cooperative attitudes in different group sizes is not straightforward. The reason is as follows: The FGF method elicits cooperative attitudes incentive-compatibly. This is because it asks for conditional and unconditional contributions from all subjects and randomly selects one of the subjects to have the conditional responses as payoff relevant; for the others in the group the unconditional contribution is used to calculate payoffs. Therefore, the probability of having the conditional response – in which we are most interested in – as payoff relevant is $1/n$ for each subject. This in unproblematic if all groups are of the same size. However, as group size increases, the incentives of the strategy method decisions change (since n increases) confounding the effect of the group size with the weights of the incentives, which become approximately hypothetical as n grows.

To overcome the problem of vanishing incentive compatibility as group size grows, we propose a new incentive-compatible strategy method to elicit conditionally cooperative preferences where the incentives for the conditional responses are independent of group size and hence scalable to any group size. The essential idea of the scalable strategy method (henceforth, SSM) is to select the unconditional and the conditional strategies with equal probability. We will introduce the details of FGF and SSM in Section 3.

In Study 1, we compare in a factorial design inspired by Isaac and Walker (1988), (i) groups sizes of 3 and 9, and (ii) MPCRs of 0.4 and 0.8. We also run all experiments using the original FGF method and SSM. Thus, Study 1 is a 2×2×2 design. Having validated the method, in Study 2 we only use the SSM and compare groups of 3 and 30 members, and with MPCRs of 0.04, 0.08, 0.4 and 0.8.

In both studies, our experiments have two stages: In Stage 1, we elicit people's conditionally cooperative preferences (either using FGF or SSM) in one of the four

MPCR×GroupSize parameterizations. This is followed, in Stage 2, by a one-shot direct response PGG (with beliefs) keeping the same parameterization. As we will explain in more detail in Section 2.3, this setting allows us to apply the ABC method, i.e., using the elicited attitudes from Stage 1 and the beliefs from Stage 2 to predict contributions in Stage 2. We then compare these predictions to the actual contributions in Stage 2.

Our main results are as follows. Consistent with previous literature (see next section) we find that beliefs and contributions increase in MPCR but not in group size. By contrast, the distribution of elicited types of cooperative preferences in Stage 1 are very similar across MPCR and group sizes and in the two methods of how to elicit them (FGF or SSM). We show that the ABC framework explains the observed contribution levels and performs at least as well with SSM than the FGF method in in-sample and out-of-sample predictions. According to the ABC approach, the MPCR effect is driven by increased beliefs, and not cooperative preferences because preferences are not much affected by game parameters. Finally, we find evidence that the effect of beliefs on contributions interacts with group size and MPCRs.

This paper is organized as follows. Section 2 describes the standard PGG and the related literature on group size, MPCR effects and the ABC of cooperation. Section 3 describes the FGF and our new scalable strategy method, SSM. Section 4 contains the details of the experimental design of Study 1. Section 5 presents the main results of Study 1. Section 6 presents the design details of Study 2 and Section 7 presents its results. Section 8 investigates the role of beliefs in more detail. Section 9 concludes.

2 Related literature, and the ABC of cooperation

2.1 The public goods game

In the linear public goods game (PGG) we use, each of n group members receives an equal endowment, e_i. Participants decide how to allocate the endowment between a private ($e_i - c_i$) and a public account (c_i). The private account has a return of 1 whereas all the money allocated to the public account (by all participants) gets multiplied by $\alpha > 1$ and the resulting amount is split equally among the n participants so that each subject gets a return of α/n, also known as MPCR, from her investment in the public account. Formally, the material payoff function for subject i is the following:

$$U_i = e_i - c_i + \frac{\alpha}{n} \sum_{j=1}^{n} c_j, \tag{1}$$

where $\frac{\alpha}{n} < 1$ in all our experiments. Assuming monetary payoff maximization, the

Nash equilibrium in this game states that all participants should contribute zero to the public account. However, the efficient allocation is achieved when all participants contribute to the public account fully. The standard result in the experimental literature is that people contribute some amounts in between those extremes (e.g., Ledyard (1995); Zelmer (2003); Gächter and Herrmann (2009)). There are different theories for why this is the case, including altruism and warm glow (Croson, 2007; Palfrey and Prisbrey, 1997; Andreoni, 1995b); confusion (Andreoni, 1995a; Houser and Kurzban, 2002; Burton-Chellew et al., 2016; Bayer et al., 2013); reciprocity (Sugden, 1984; Weber et al., 2018; Isler et al., 2021); matching behavior (Guttman, 1986); inequality aversion (Fehr and Schmidt, 1999); and guilt aversion (Chang et al., 2011; Dufwenberg et al., 2011). See also Katuščák and Miklánek (2022) for a recent comparative analysis. In the following, we review the most relevant literature for our purposes, that is, the role of MPCR and group size for cooperation and the ABC approach to explaining cooperation.

2.2 MPCR and group size effects

MPCR effects - the higher the marginal per capita return in a public goods game, the higher the level of cooperation - have long been observed in the literature (see, e.g., Ledyard (1995); Palfrey and Prisbrey (1997); Brandts and Schram (2001); Goeree et al. (2002) and Zelmer (2003)). MPCR effects are also closely linked to the question of how group size affects voluntary cooperation. As can be seen from the payoff function shown in Eq. 1, if n increases, the MPCR (α/n) also diminishes. To isolate a pure group size effect therefore requires controlling for MPCR, which is what Isaac and Walker (1988) achieved in their seminal experiments.

Isaac and Walker (1988) found that group size increased cooperation but only if the MPCR changed with group size. When the MPCR was held constant they found no group size effects. Since then, a number of papers have studied the effect of group size on contributions to PGGs.[2] For instance, Pereda et al. (2019a) found that groups of 100 and 1000 are not significantly different in terms of cooperation which suggests that extrapolating conclusions from smaller group sizes might be reasonable. In a within-subjects design, with group sizes of 5, 10, ..., 40, Pereda et al. (2019b) observed a monotonic increase of the cooperation rate as a function of group size.

Diederich et al. (2016) also found a positive group size effect and concluded that the effect was driven by the intensive margin (i.e., those who contribute contributed more in larger group sizes). This is consistent with the beliefs of the subjects. They also show that the share of free riders remains constant in spite of the beliefs of the

[2] Further experiments on MPCR and group size include Marwell and Ames (1979); Isaac et al. (1994); Goeree et al. (2002); Carpenter (2007); see also Zelmer (2003) and Ledyard (1995) for overviews.

participants showing a larger share of free riders in larger groups. An older study (Kerr, 1989) argued that beliefs are also affected by an efficiency illusion (i.e., a diminished sense of self-efficacy in larger groups) even when the self-efficiency was kept constant.

Nosenzo et al. (2015) and Weimann et al. (2019) showed that the effect of group size depends on the MPCR level. When the MPCR is low, the effect of group size on contributions is positive, but the effect disappears or becomes negative when the MPCR is high. Weimann et al. (2019) argue that the reason is due to beliefs, which in turn are affected by the saliency of the advantages of cooperation. They propose a measure of this saliency and show the change in this measure with respect to group size is higher in PGGs with low MPCRs.

In summary, beliefs take a central role in most recent studies, but how conditionally cooperative preferences are influenced by group size and MPCR remains unexplored. Moreover, it is unknown whether the ABC approach works for larger group sizes and high MPCRs. The aim of this paper is to provide the missing evidence.

2.3 The ABC of cooperation

Our ambition in this paper is not only to provide evidence on conditionally cooperative preferences as a function of group size and MPCR but also to explore the explanatory power of conditionally cooperative preferences jointly with beliefs to explain cooperation levels across a variety of PGG parameter sets. The basic idea was provided by Fischbacher and Gächter (2010) who used the strategy method to explain the decline of cooperation. They elicited people's preferences for conditional cooperation (details follow in the next section) and had subjects play ten rounds of a repeated public goods game played in the Stranger matching protocol (that is, group composition is changed at random in each round). In each round, players made a contribution decision and also reported their beliefs about how much their current group members would contribute in the next round. Fischbacher and Gächter (2010) then showed that the elicited cooperative preferences evaluated at the beliefs of a given period predicted actual contributions round by round. Fischbacher et al. (2012) further showed that this method also predicts contributions well in one-shot public goods games.

Gächter et al. (2017) named this method the ABC approach to cooperation. It measures individual *a*ttitudes (a_i) to cooperation as a function of all possible rounded average contributions of other group members (sometimes also called dispositions or preferences), *b*eliefs (b_i) about others' contributions average contributions and actual *c*ontributions (c_i) separately and explains cooperation as $a_i(b_i) \rightarrow c_i$. They found that the ABC approach explains contributions in maintenance and provision versions of public goods (see also Gächter et al. (2022) for a related result).

Interestingly, Gächter et al. (2017) and Gächter et al. (2022) observed that the share of conditional cooperators was higher (and the share of free riders lower) in the provision version of the public good game than in the maintenance version, which suggests that attitudes to cooperation are influenced by the context of the game. Isler et al. (2021) replicated these findings and present a framework that generalizes the ABC approach to what they call the Contextualized Strong Reciprocity Approach (CSR).[3] CSR is a version of the ABC approach that allows for context effects: CSR explains cooperation as $a_i(f, b_i(f)) \rightarrow c_i$, where both a_i and b_i are potentially functions of the contextual features (f) of a given public goods game. In this paper, f is the MPCR or the group size, both of which were very similar in all previous studies that used the ABC approach (in Isler et al. (2021), f referred to maintenance or provision public good). Our goal in this paper is to provide evidence on all elements of $a_i(f, b_i(f)) \rightarrow c_i$ where the experiments will manipulate f as group size and MPCR.

3 Strategy methods

The strategy method was introduced by Selten (1967) in the context of an oligopoly game but has now been used in all kinds of games (Brandts and Charness (2011); Keser and Kliemt (2021)). Here we focus on the application of the strategy method to PGG (for a review of recent literature, see Thöni and Volk (2018)). The idea is to ask subjects to make contingent decisions for each possible average contribution of other members in the group. By eliciting a complete contribution profile, we get data on all possible situations including those that are only rarely reached (e.g., everyone in the group contributing zero or fully). We will first review the standard method by Fischbacher et al. (2001) and present its problems in the context of our research questions in Section 3.1 and then introduce our new version, called the scalable strategy method, in Section 3.2.

3.1 Fischbacher, Gächter and Fehr (2001; FGF)

The first application of the strategy method in the one-shot public goods game is due to Fischbacher et al. (2001) (FGF). In the FGF method, participants are asked to make two decisions: (i) an unconditional contribution and, (ii) a contribution table. The unconditional contribution is the amount they would like to contribute without any information on the contribution of the other group members. For the contribution table, participants need to answer how much they would like to contribute for each

[3] Conditional cooperation is one important instance of "strong reciprocity". See Weber et al. (2018) and Isler et al. (2021) for discussions and analyses.

possible average contribution of the other group members (i.e., the contributions they would like to make conditional on the other group member's contributions as if they were last movers in a sequential game). Participants are told that after the decisions are made, $n-1$ of them will be selected randomly to have the unconditional contribution used as their payoff relevant decision and the remaining group member will have their contribution table used as their payoff relevant decision (computed by imputing the average of the other $n-1$ members in their contribution table). See the Online Appendix for the instructions.

The FGF method is an incentive-compatible way to elicit the conditional contribution decisions (since all the choices are potentially payoff relevant). For our research purposes, it is important to note that the probability of a participant having the contribution table selected as payoff relevant depends on the number of participants in the group (i.e., $\frac{1}{n}\%$). Therefore, the larger the group, the more hypothetical the contribution table becomes. This posits a problem if we want to incentive-compatibly elicit cooperative dispositions as a function of group size. In the following subsection, we introduce a method that solves this problem.

3.2 A scalable strategy method (SSM)

Our research questions demand a method that keeps the probability of the contribution table being payoff relevant constant ("scalable"). We achieve this by giving the contribution table and the unconditional contribution equal weights (50% each). So, instead of selecting one participant at random, we randomize for each participant whether they would have the unconditional contribution or the conditional table as payoff relevant. In this way, the probability of playing with the conditional table is 50% for each player in any group size and we achieve both goals. Because both decisions, the unconditional contribution and the contribution table, are payoff relevant, our elicitation of the contribution table is incentive compatible.

The SSM procedure involves two steps: First, we calculate the *computed conditional contribution* for player i by plugging the average contribution of the remaining players into player i's table. Second, for each player i, we randomly select either the unconditional contribution or the contribution table with 50% probability each. If the unconditional contribution is chosen, we use all other group members' computed conditional contributions to determine i's individual payoff; otherwise, if the contribution table is chosen, we use the computed conditional contribution for player i and all other player's unconditional contributions to determine i's individual payoff (see the Appendix of this paper for further details and a concrete example of the procedure).

Notice that it is perfectly possible to have groups where either the unconditional contribution or the contribution table is selected for all players. By contrast, in the FGF method, we cannot have all players to have their payoff determined by their conditional contribution (as indicated in their payoff table) because we need some value to plug into the other players' table. In our case, we pin down the value from the table in the first step of the procedure.

SSM comes at a cost: the size of the public good is individual-specific, as opposed to the standard PGG where the same size of the public good applies to everyone in the group. However, there are at least three reasons why this might not a be a major concern for our purposes. First, we are still studying a social dilemma (i.e., the Nash equilibrium is to contribute zero, and efficiency demands to contribute everything). Second, this problem is independent of group size or MPCR. Third, and most importantly, we are only interested in the elicited conditional preferences and not the unconditional contributions. The conditional preferences are elicited incentive-compatibly and the incentives for their truthful revelation are independent of group size.

4 Study 1: Design

The experiments consisted of a 2×2×2 between-subjects design varying the method (FGF and SSM), group size (small = 3 and large = 9) and the MPCR (low = 0.4 and high = 0.8). Participants (n = 936) were randomized into one of the eight conditions displayed in Table 1. Each group member had an endowment of 10 tokens they could either keep or invest into the public good (see Section 2.1). The strategy methods therefore elicited contributions for each of the eleven (0 to 10) possible average contributions of other group members.

All experiments had two stages. In Stage 1, participants played the PGG in either the FGF or SSM strategy method. In Stage 2, they played the PGG in the direct-response mode. Apart from different parameters in the respective condition (see Table 1) the Stage 2 PGGs were the same regardless of the strategy method used in Stage 1.

After explaining the game, we used three incentivized control questions: two questions about the PGG and one regarding the probability of having the contribution table as the payoff relevant decision (see the Online Appendix for instructions and control questions). Participants who responded incorrectly in the first attempt had to keep trying until they got it right before proceeding (but only those responding correctly in the first attempt were compensated for their correct answers). Participants then completed the respective strategy method task (FGF or SSM, depending on the condition they were randomly selected into). This completed Stage 1 of the experiment.

In Stage 2, participants made a contribution decision in a one-shot direct response

Condition	Stage 1 Strat. method	MPCR	Group size (n)	Social return (α)	N
1	FGF	0.4	3	1.2	117
2	FGF	0.4	9	3.6	117
3	FGF	0.8	3	2.4	117
4	FGF	0.8	9	7.2	117
5	SSM	0.4	3	1.2	117
6	SSM	0.4	9	3.6	117
7	SSM	0.8	3	2.4	117
8	SSM	0.8	9	7.2	117

Table 1: Study 1 experimental conditions (between-subjects)

PGG with the same parameters as relevant in the respective condition (see Table 1). We also elicited the beliefs (first and second order). We incentivized the belief elicitation by rewarding subjects if their prediction was exactly right. We will use these games to test for the predictions made by the ABC approach (see Section 2.3).

The experiment was programmed in Qualtrics and conducted online using the platform Prolific with participants from the UK in the 18-25 years old range.[4] We restrict the participant sample by age in order to minimize potential confounds caused by age and cohort effects.[5] Study 1 was run between January and February 2021. Participants were paid the equivalent to £12.71 per hour on average. The average participant took around 8 minutes to complete the experiment.

5 Study 1: Results

5.1 Attitudes

Both strategy methods allow us to classify participants into types according to their cooperative preferences ("attitudes" a_i). Here we follow Thöni and Volk (2018) and classify the attitude types as Conditional Cooperators, Triangle Cooperators, Unconditional Cooperators and Free Riders.[6]

[4]Hergueux and Jacquemet (2015) show that the online elicitation of social preferences and behaviour is reliable. See also Arechar et al. (2018) for a discussion of online experiments. Palan and Schitter (2018) discuss the use of Prolific for online experiments.

[5]Age is known to influence cooperation positively, see, e.g., List (2004); Gächter and Herrmann (2011); Arechar et al. (2018). To maximize the statistical power of our study, we kept the age of our participants constant and similar to the vast majority of studies on MPCR and group size effects.

[6]For this we used the Stata command "cctype" provided in Thöni and Volk (2018). In words, Conditional Cooperators increase their cooperation with the cooperation of others, Triangle Cooperators increase their contributions up to a certain point and then decrease with the contribution of others, Unconditional Cooperators contribute a non-zero constant value for any contribution of others, and Free Riders contribute zero for any contribution of others.

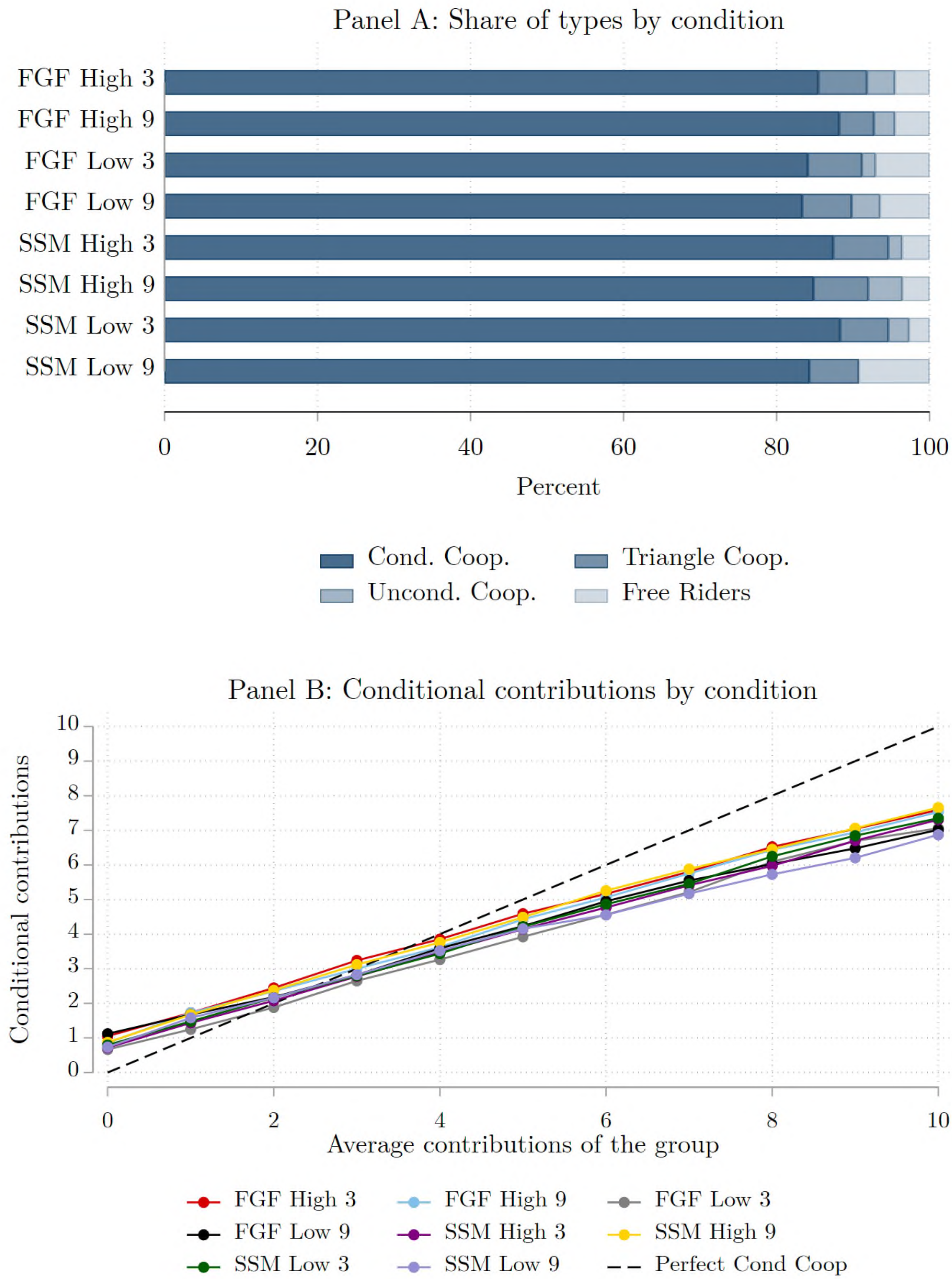

Figure 1: Cooperative attitudes a_i as elicited in Stage 1 in Study 1. Panel A: Distribution of conditionally cooperative types. Panel B: Average conditional contribution $\bar{a}_i$ across all preference types. High (Low) refers to an MPCR of 0.8 (0.4). Large (Small) refers to a group size of 9 (3).

Figure 1A shows that the distribution of types is remarkably similar in all eight conditions. Conditional Cooperators is the main category, which accounts for 80.9% of the subjects, followed by the Triangle Cooperators 6.1% and Free Riders with 4.9%. We also compare the attitude type distributions between FGF and SSM for each of the four parameterizations using Chi-squared tests and find that none of the pairwise comparisons is significant (smallest p-value = 0.338). This suggests that compared to FGF, SSM is equally good in identifying attitude types.

Figure 1B, shows the average of the responses in the contribution tables (i.e., the average conditional contributions $\bar{a}_i$). As a benchmark, we also plot the line for perfect conditional cooperation (the diagonal). Average conditional cooperation $\bar{a}_i$ is almost the same in all conditions (i.e., by method, MPCR, and group size).

	(1) SSM Cond. contrib.	(2) FGF Cond. contrib.
large	-0.193	0.216
	(0.232)	(0.252)
high MPCR	-0.078	0.524**
	(0.226)	(0.244)
large x high MPCR	0.540	-0.336
	(0.333)	(0.351)
_cons	4.149***	3.932***
	(0.160)	(0.176)
R^2	0.009	0.011
N	468	468

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable is the average of all conditional contributions in either FGF or SSM. Large is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. High MPCR is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 2: Cooperative attitudes a_i as elicited in Stage 1 by SSM and FGF in Study 1 as a function of group size and MPCR

To test the link between game parameters and conditional cooperation formally, we averaged the conditional contributions for each subject and regress them on dummies for high MPCR, large group and their interaction. The results are shown in Table 2.[7]

[7]The results are virtually unchanged if we use all 11 responses and cluster the standard errors by subject.

Although the results are, with one exception, not statistically significant, we see differences between the methods. In Column 1 (SSM), we see that conditional cooperation is insignificantly lower in the large group and in MPCR, but the effect reverses insignificantly when the MPCR is high. In Column 2 (FGF), we see something different, that is, conditional cooperation increases insignificantly with group size and significantly with MPCR but decreases insignificantly with the interaction of the two.

5.2 Beliefs and contributions

After the strategy method PGG experiment in Stage 1, subjects in Stage 2 participated in a one-shot direct-response PGG where we observed c_i and where we also elicited beliefs b_i. Recall that these one-shot PGG games in Stage 2 had the same parameter conditions as in Stage 1, regardless of whether we used FGF or SSM in Stage 1. Figure 2 shows the averages of beliefs and contributions by treatment and suggests a positive MPCR effect in both beliefs and contributions, but no group size effect.

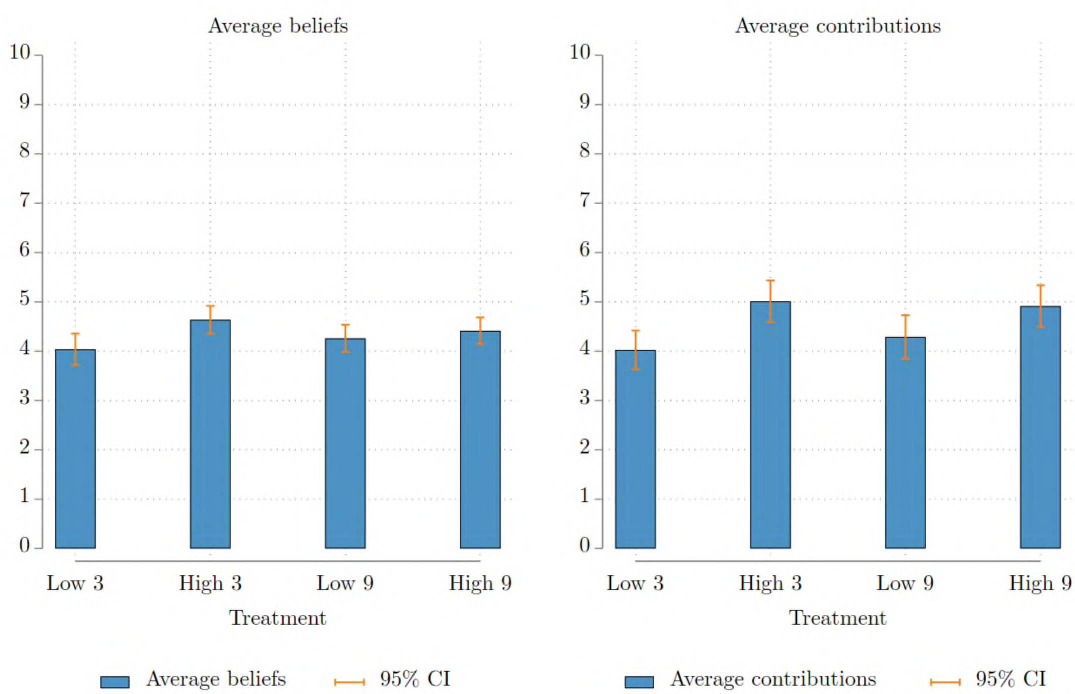

Figure 2: Average of beliefs b_i and contributions c_i by treatment in Stage 2 of Study 1. High (Low) refers to an MPCR of 0.8 (0.4). Large (small) refers to group size of 9 (3).

In order to study the differences between the treatments formally, we regress beliefs and contributions on dummy variables for large group ($n = 9$) and high MPCR (0.8). For this, we pool the observations from both strategy methods since, for given game parameters, the direct response one-shot public good game in Stage 2 is the same regardless of the strategy method used in Stage 1. Table 3 presents the results. We find positive and significant effects of the MPCR on beliefs and contributions. These results confirm the MPCR effects reported in previous literature (see Section 2.2). Regarding group size, we find a positive effect but it is not statistically significant. Figure A.1 in the Online Appendix presents the full distributions.[8]

	(1) Pooled beliefs	(2) Pooled contrib.
large	0.222	0.265
	(0.213)	(0.300)
high MPCR	0.598***	0.987***
	(0.215)	(0.293)
large x high MPCR	-0.444	-0.363
	(0.291)	(0.426)
_cons	4.038***	4.026***
	(0.161)	(0.200)
R^2	0.010	0.016
N	936	936

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variables are the beliefs (column 1) and contributions (column 2) reported in the Stage 2 direct-response one-shot public good game. Large is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. High MPCR is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 3: Beliefs and contributions in Stage 2 of Study 1 as a function of group size and MPCR

5.3 Testing the accuracy of ABC predictions

Having shown the descriptive evidence, we proceed by testing the predictive accuracy of the ABC approach. We start by using the elicited beliefs b_i and check how well they

[8]For completeness, Figure A.2 in the Online Appendix presents the distributions of the second-order beliefs. However, for the ABC approach only first-order beliefs matter, which is why we focus on them here.

can predict actual contributions c_i by using the attitudes a_i as described in subsection 2.3. We compute the prediction error as the difference between the actual contributions and the contributions predicted by the ABC approach.

Formally, the predicted contribution of individual i is calculated as $\hat{c}_i = a_i(b_i)$; the prediction error therefore is $\hat{c}_i - c_i$. Individual $i's$ $\hat{c}_i$ can be derived non-parametrically (by taking $i's$ contribution c_i as indicated in $i's$ table of conditional contributions a_i at the belief b_i) or parametrically (by estimating the slope of the conditional contributions $\hat{a}_i$ from all entries in individual $i's$ vector a_i and using these parameters and the beliefs to calculate $\hat{c}_i = \hat{a}_i(b_i)$). Since the large majority of subjects are (imperfect) conditional cooperators, we follow the second approach to obtain smoother estimates, but the results are qualitatively similar if we follow the non-parametric approach (see Online Appendix Section B).

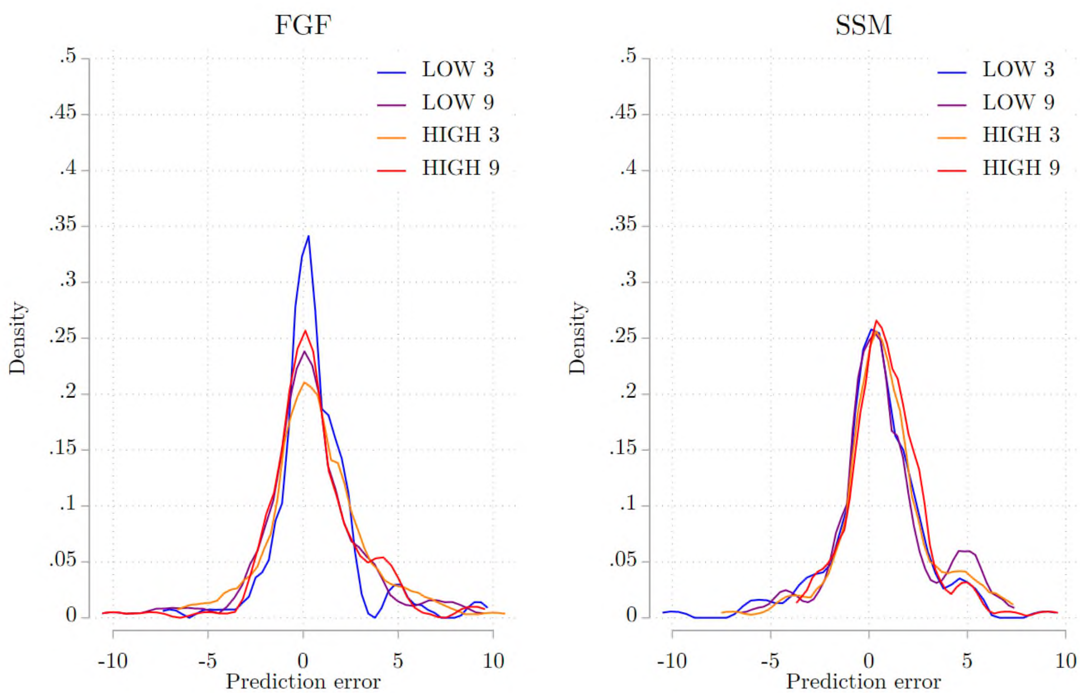

Figure 3: Distribution of ABC prediction errors ($\hat{c}_i - c_i$) in Study 1

Figure 3 shows that in all conditions the mode of the $\hat{c}_i - c_i$ distributions is at zero, which means actual contributions coincide with the contributions predicted by the ABC approach. Figure 3 also suggests that the distributions are similar across conditions. Note, however, that the distributions in different treatments overlap slightly more under SSM. Kolmogorov-Smirnoff tests comparing the prediction errors under FGF and SSM in each treatment only reject the null hypothesis of equal distributions in the High 9 condition (p-value=0.046), but this does not survive a Bonferroni correction for multiple comparisons (see the Online Appendix Figure A.4 for treatment-specific graphs).

	(1) Pooled FGF	(2) Pooled SSM	(3) Large (9) FGF	(4) Large (9) SSM	(5) Small FGF	(6) Small SSM
predicted	0.555***	0.615***	0.521***	0.670***	0.578***	0.560***
	(0.0919)	(0.0676)	(0.140)	(0.0911)	(0.124)	(0.101)
beliefs	0.405***	0.407***	0.467**	0.486***	0.362**	0.344***
	(0.118)	(0.0807)	(0.193)	(0.0996)	(0.152)	(0.124)
high MPCR	0.222	0.393*	0.223	0.268	0.233	0.510*
	(0.226)	(0.209)	(0.331)	(0.283)	(0.305)	(0.302)
large	-0.174	0.269				
	(0.220)	(0.202)				
_cons	0.528*	0.127	0.222	-0.108	0.614**	0.567
	(0.279)	(0.250)	(0.455)	(0.269)	(0.308)	(0.367)
R^2	0.489	0.547	0.454	0.615	0.528	0.484
$CVMSE$	**5.938**	**4.807**	**6.650**	**4.306**	**5.482**	**5.303**
N	468	468	234	234	234	234

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the contributions in the one shot public good game. The beliefs are the ones reported in the one shot public good game and predicted is the prediction using the either FGF or SSM. Large is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. High MPCR is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 4: Regressions explaining contributions in Study 1 using the ABC method

Table 4 shows the regressions of the basic ABC model used in Fischbacher and Gächter (2010) and Gächter et al. (2017) with the contributions c_i in the one-shot PGG as dependent variable. This approach includes predicted contributions $\hat{c}_i$ and beliefs b_i as additional regressors because Fischbacher and Gächter (2010) and Gächter et al. (2017) have shown that $c_i = \beta\hat{c}_i + (1 - \beta)b_i$, that is, actual conntributions are a weighted average of predicted contributions and beliefs. In other words, beliefs matter on top of predicted contributions (presumably because of the enhanced saliency of a specific belief in a direct response PGG). In addition we add a dummy for high MPCR and a dummy for large group size (in the pooled models of Columns 1 and 2). Columns 1, 3 and 5 show the FGF method and columns 2, 4 and 6 the SSM for different samples (full sample, large, small).

The estimation results confirm previous findings that both beliefs and predicted contributions are highly significantly positively correlated with contributions in all models. High MPCR does not contribute significantly on top of beliefs and predicted contributions (except in Columns 2 and 6 where we find a weakly significantly positive effect).

Our next step is to compare the two methods in terms of in-sample and out-of-sample prediction power (as in, e.g., Andreoni et al. (2015)). The cross-validated Mean Squared Error (CVMSE) is a measure of out-of-sample prediction, in other words, how well the estimated model predicts the contributions of new observations. To compute this we split the sample randomly in five sub-samples and train the data (i.e., estimate the parameters) using four of them to predict the remaining one.[9] After that, we compute the mean squared error of these predictions.

For the full sample comparison, we find that the SSM outperforms FGF because the CVMSE is smaller in SSM than in FGF (4.807 vs. 5.938). We should also expect this relative increase in performance to be higher when the group size is larger. This is because the probability of the contribution table being payoff relevant changes with group size in FGF but not in the SSM. In the experiment, the probabilities are 0.33 (FGF) vs. 0.50 (SSM) in groups of 3 members and 0.11 (FGF) vs. 0.50 (SSM) in groups of 9 members (see Section 3.1). To see this, we split the data by group sizes in columns 3 to 6. The CVMSE in the large groups is 6.650 in FGF and 4.306 in SSM. Even in the small group, SSM slightly outperforms FGF in out-of-sample predictions (5.303 vs. 5.482). Similarly, the R-squared (in-sample prediction) is 35% higher in the large group and 12% higher in the full sample, but 8% lower in the small group.[10] The main results are also robust to restricting the sample to conditional cooperator types only and to including a dummy for part-time employed (i.e., the only variable that was unbalanced between FGF and the SSM).

In summary of Study 1, we find that neither MPCR nor group size affect conditional preferences a_i; beliefs b_i and contributions c_i do significantly increase in MPCR but not group size, confirming previous literature (see Section 2.2). Because attitudes are not affected by game parameters, but MPCR increases beliefs, the ABC approach predicts higher contributions when the MPCR is high, and our experiments confirm this prediction.

While our results are reassuring, some caution is warranted: although our large group was three times larger than the small group, we are still in a relatively small group size context. Next, we therefore explore significantly larger group sizes.

[9] We used the Stata command "crossvalidate" (Schonlau, 2020).

[10] It is also possible that the SSM outperforms FGF because it is simpler to understand (i.e. 50% probability of each decision being payoff relevant as opposed to picking a participant at random with an ad-hoc probability). However, why this is the case is beyond the scope of this paper.

6 Study 2: Design

For Study 2, we compare group sizes of 3 and 30 (a ten-fold increase). This poses a new problem. Keeping the MPCR constant would imply a huge social return α. Thus, instead of holding the MPCR (α/n) constant, we hold α constant. The parameters are shown in Table 5. Because we found in Study 1 that the SSM performs well, especially in large samples, we rely on it for Study 2

In summary, the experiment in Study 2 consisted of a 2×2 between-subjects design using the SSM only, varying group size (small = 3 and large = 30) and α (low = 1.2 and high = 2.4). Apart from these differences, the Study 2 experiment was very similar to the Study 1 experiment: as in Study 1, Study 2 comprised two stages. In Stage 1 we elicited attitudes to cooperation using the SSM method, and in Study 2 participants played a direct response PGG (with the parameterization of the respective condition). We also elicited first- and second-order beliefs.

Study 2 was run in June 2021 with 600 participants on Prolific who were paid the equivalent to £11.93 per hour on average. The average subject took around 8 minutes to complete the experiment. Study 1 participants could not take part in Study 2.

Condition	Group size (n)	Social return (α)	MPCR	N
1	3	1.2	0.4	150
2	30	1.2	0.04	150
3	3	2.4	0.8	150
4	30	2.4	0.08	150

Table 5: Study 2 experimental conditions (between-subjects) using SSM in Stage 1

7 Study 2: Results

7.1 Attitudes

Figure 4A reveals that the share of types is very similar in all conditions. In the low α treatments, nonetheless, there is some evidence that the number of free riders increases (from 2% to 6.67%) and the number of conditional cooperators decreases (from 82.67% to 78.67%) with group size. This also happened in Study 1 (SSM) for the Low MPCR treatments, the number of free riders increases (from 2.56% to 8.55%) and the number of conditional cooperators decreases (from 83.76% to 77.78%) with group size.

Nosenzo et al. (2015) and Weimann et al. (2019) found that, in low MPCR treatments, cooperation increases with group size, but they do not disentangle the role of

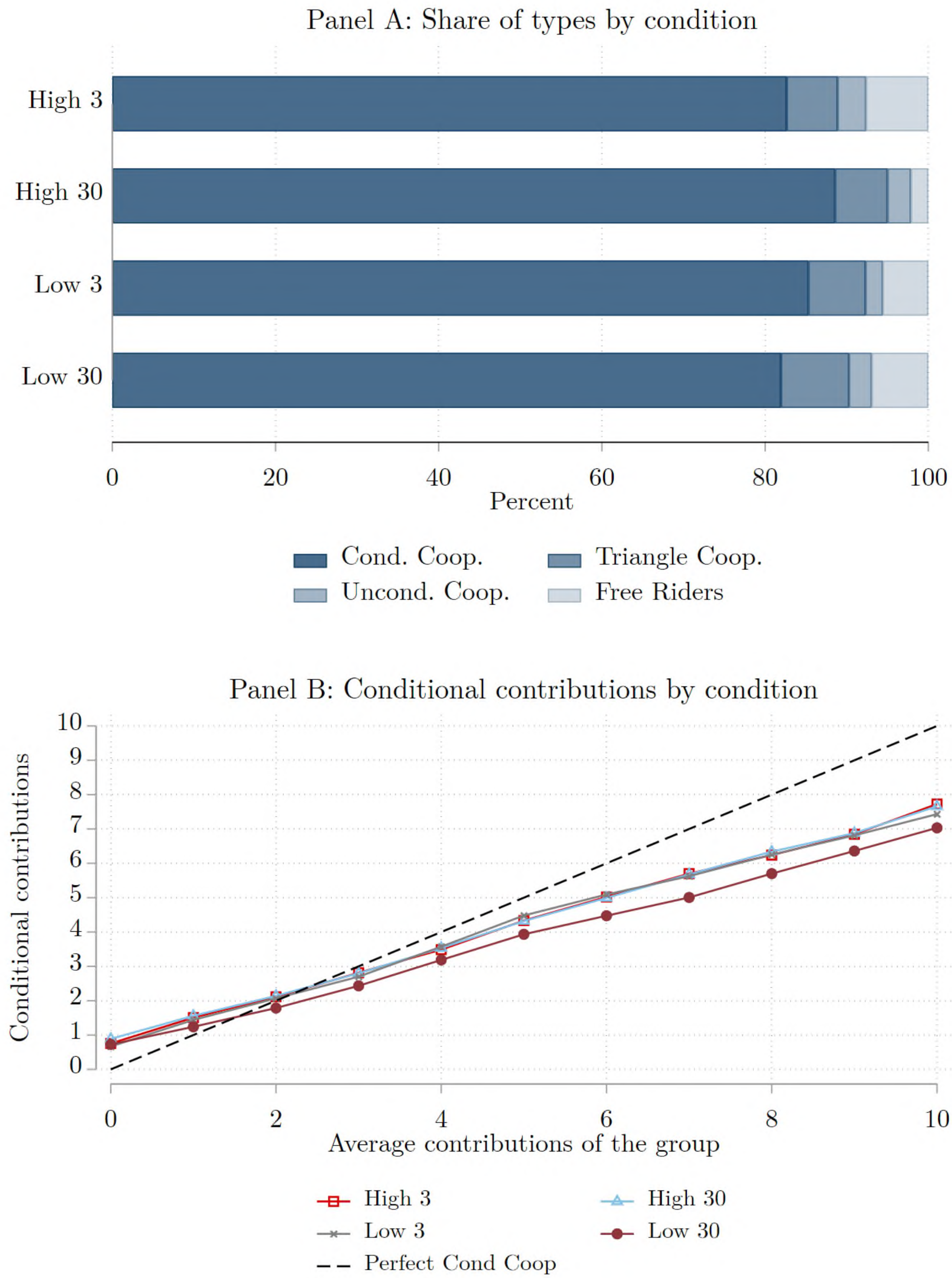

Figure 4: Cooperative attitudes a_i as elicited in Stage 1 of Study 2 (SSM only). Panel A: Distribution of conditionally cooperative types. Panel B: Average conditional contributions $\bar{a}_i$ across all preference types. High (Low) refers to a social return of 2.4 (1.2). Large (small) refers to group size of 30 (3).

conditional preferences and beliefs for cooperation. Our results would be consistent with theirs if conditional cooperators in larger groups have a slope high enough to compensate for the increase in free riders and the reduction in conditional cooperators. This argument is consistent with Diederich et al. (2016), who found a positive group size effect driven by the intensive margin (see Section 2.2).

Figure 4B shows the average of the responses in the contribution tables by condition (with the dashed black line representing a hypothetical perfect conditional cooperator). Again, the slopes follow a similar pattern in all four conditions.

	(1) Cond. contrib.	(2) beliefs	(3) contrib.
large	-0.447**	-0.207	-0.207
	(0.211)	(0.242)	(0.354)
high α	-0.023	0.693***	0.560
	(0.205)	(0.264)	(0.369)
large x high α	0.411	-0.220	0.187
	(0.299)	(0.355)	(0.531)
_cons	4.254***	4.053***	4.207***
	(0.134)	(0.178)	(0.240)
R^2	0.010	0.023	0.011
N	600	600	600

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. SSM only in Stage 1. The dependent variable is the average of all conditional contributions a_i as elicited in Stage 1 of Study 2 (model 1), the beliefs b_i (model 2) and contributions c_i (model 3) as elicited in the direct-response one-shot public good game in Stage 2 of Study 2. Large is an indicator that takes the value 1 for groups of 30 and 0 for groups of 3. High α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Robust standard errors in parentheses.

Table 6: Cooperative attitudes a_i; beliefs b_i; and contributions c_i, all as a function of group size, MPCR, and their interaction

In Table 6 we test formally how game parameters affect the attitudes to cooperation. Column 1 shows that preferences for cooperation are lower in the large groups, but the interaction term with high α is positive and of the same magnitude, meaning that only large groups with low α have significantly lower preferences for cooperation. We see that the groups of 3 with MPCR=0.4 (α = 1.2) and MPCR=0.8 (α = 2.4) are more similar than the groups of 30 with MPCR=0.04 ($\alpha = 1.2$) and MPCR=0.08 ($\alpha = 2.4$). We interpret this as group size increasing the effect of α on preferences for cooperation, even when the MPCR change is much smaller in absolute terms.

7.2 Beliefs, contributions, and ABC prediction errors

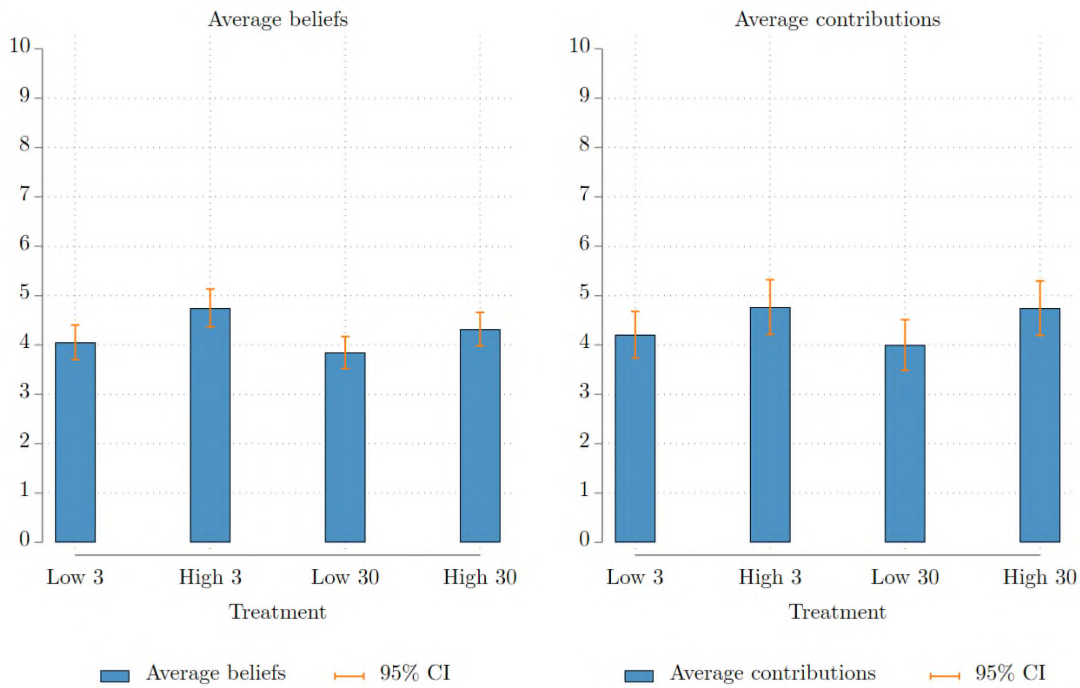

Figure 5: Average of beliefs b_i and contributions c_i by treatment in Stage 2 of Study 2. High (Low) refers to social returns of 2.4 (1.2). Large (small) refers to group size of 30 (3).

Figure 5 presents the average beliefs b_i (left) and contributions c_i (right) in each treatment of Study 2. Consistent with the results from Study 1 (see Fig. 2), beliefs and contributions increase in the social return (α) and suggest no group size effect. The formal analysis is presented in Columns 2 and 3 of Table 6, which shows the regressions of beliefs and contributions on dummy variables for groups of 30 and $\alpha = 2.4$. We find a positive effect and similar in magnitude for α in both cases but only statistically significant for beliefs. Again, we find no statistically significant effect of group size in beliefs or contributions.[11] Like in Study 1, the prediction errors are similar across conditions, as shown in Figure 6. Overall, the results are consistent with Study 1.

[11]In the Online Appendix Figure A.5, we present the distribution of beliefs and contributions holding constant the group size or the MPCR, separately. Figure A.6 presents the distributions of the second order beliefs.

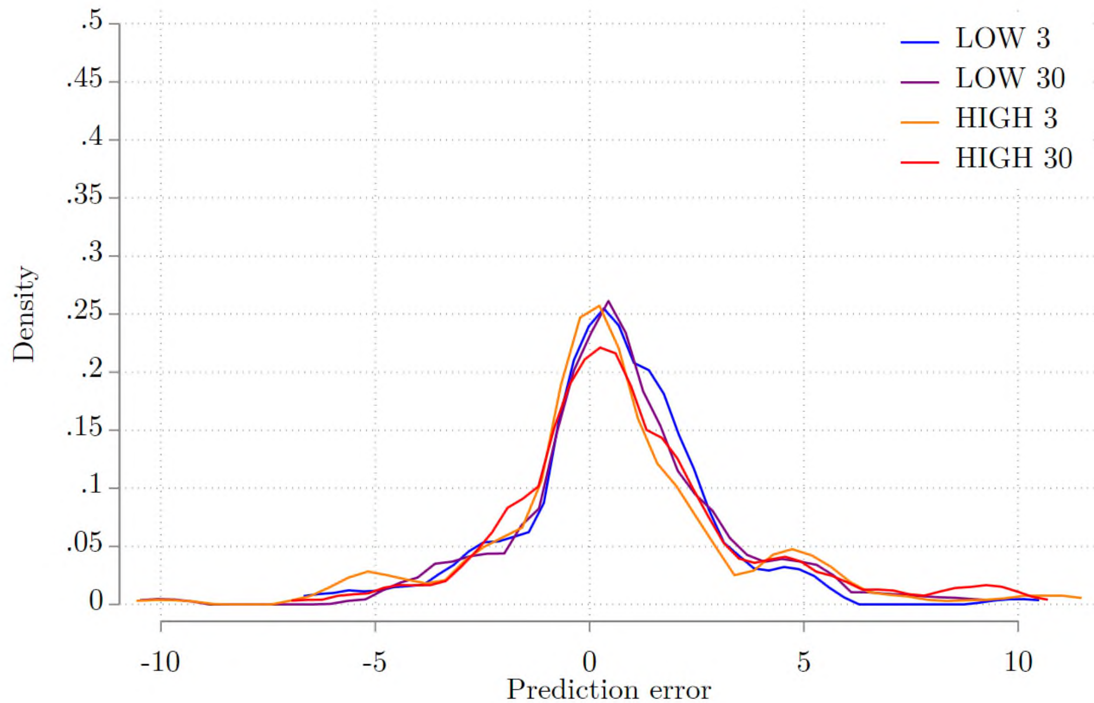

Figure 6: Distribution of ABC prediction errors ($\hat{c}_i - c_i$) in Study 2

Finally, Table 7 shows the regressions of the basic ABC model. Although we are holding α instead of the MPCR constant, the results are remarkably similar to Study 1, the coefficients of *beliefs* and *predicted* are economically and statistically significant and the coefficients of large and high α small and non-significant.

8 MPCR, group size, and the role of beliefs for cooperation

Confirming previous results in the literature (see Section 2.2), we find that, on average, MPCR has a positive effect on contributions (Tables 3 and 6). Regarding group size the same tables show no clear pattern of the average effect on contributions (positive in Study 1 and negative in Study 2).

The CSR approach (see Section 2.3) proposes that a_i and b_i can be functions of contextual features f (i.e., MPCR and group size). In theory, the CSR approach should internalize f in *predicted*, but in practice Fischbacher and Gächter (2010) found that *beliefs* matter on top of the predictions to explain contributions - an observation we also made in Study 1 (Table 4) and Study 2 (Table 7). Although with large standard

	(1) All	(2) Large (30)	(3) Small
predicted	0.571***	0.586***	0.552***
	(0.080)	(0.106)	(0.121)
beliefs	0.329***	0.379***	0.297**
	(0.101)	(0.139)	(0.146)
high α	0.147	0.244	0.052
	(0.204)	(0.300)	(0.276)
large	0.217		
	0.203		
_cons	0.713***	0.627*	0.977***
	(0.267)	(0.342)	(0.359)
R^2	0.421	0.430	0.415
$CVMSE$	**6.356**	**6.539**	**6.278**
N	600	300	300

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the contributions in the Stage 2 one-shot direct-response public good game. Large is an indicator that takes the value 1 for groups of 30 and 0 for groups of 3. High α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Column 1 uses the full sample and columns 2 and 3 split the sample by group size. CVMSE is the cross validated mean square error. Robust standard errors in parentheses.

Table 7: Regressions explaining Stage 2 contributions in Study 2 using the ABC approach

errors, Table 4 and Table 7 also show large coefficients for MPCR and group size. Here we explore whether f governs the effect of beliefs on contributions (e.g., beliefs might become a more important component in the ABC regressions as MPCR or group size increases). In order to test this formally, we re-run, for each study, the ABC regressions including beliefs interacted with group size and MPCR (or α).

Table 8 presents the results. Column 1 shows the results for Study 1 using the SSM. We find a strong negative effect of group size on contributions. The interaction of *beliefs* x *large* indicates that the effect size depends on the level of beliefs. Mathematically, $\frac{\partial c}{\partial \text{large}} = -0.803 + 0.245 * \text{beliefs}$. The direct negative effect gets reversed if beliefs are high enough, in our sample, that is $0.803/0.245 = 3.28$. These results are not replicated in the FGF sample (Column 2) but they survive in the pooled sample (Column 3).

	(1) Study 1 SSM contrib.	(2) Study 1 FGF contrib.	(3) Study 1 Pooled contrib.	(4) Study 2 SSM contrib.
predicted	0.601***	0.550***	0.570***	0.571***
	(0.069)	(0.092)	(0.0582)	(0.080)
beliefs	0.189*	0.364***	0.288***	0.263**
	(0.114)	(0.126)	(0.0851)	(0.116)
large	-0.803**	-0.345	-0.542	-0.286
	(0.420)	(0.523)	(0.338)	(0.474)
beliefs x large	0.245***	0.043	0.139*	0.119
	(0.093)	(0.122)	(0.0776)	(0.112)
high MPCR	-0.636	-0.036	-0.316	
	(0.433)	(0.523)	(0.342)	
beliefs x high MPCR	0.236**	0.062	0.147*	
	(0.099)	(0.121)	(0.0786)	
high α				0.049
				(0.478)
beliefs x high α				0.024
				(0.110)
_cons	1.123	0.706	0.872***	0.996***
	(0.426)	(0.347)	(0.270)	(0.380)
R^2	0.560	0.490	0.520	0.422
N	468	468	936	600

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the contributions in the one shot public good game. The beliefs are the ones reported in the one shot public good game and predicted is the prediction using the either FGF or SSM. Large is an indicator that takes the value 1 for groups of 9 or 30 and 0 for groups of 3. High MPCR is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. High α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Robust standard errors in parentheses.

Table 8: The role of beliefs in explaining cooperation, controlling for ABC predictions, group size, and MPCR

Regarding MPCR, we find a strong negative effect but with large standard errors. The interaction term *beliefs* x *MPCR* is positive, statistically significant and similar in magnitude to the interaction with group size. Mathematically, $\frac{\partial c}{\partial \text{MPCR}} = -0.636 +$

$0.236 * \text{high}$. As in the group size case, the negative effect gets reversed when beliefs are high enough, here the threshold is $0.636/0.236 = 2.69$. Again, we do not reject any of these hypothesis in the FGF method but the results survive in the pooled sample.

Interestingly, we find that the beliefs threshold to make the overall group size effect positive (3.28) is higher than the beliefs threshold to make the overall MPCR effect positive (2.69). This could explain why there is more consensus in the effect of MPCR than on the group size effects in the literature.

Column 4 shows the regression results for Study 2. Although the group size effect is not statistically significant, the results are qualitatively similar. We conjecture that this is due to the MPCR adjustment (i.e., holding the α constant instead of the MPCR) offsetting the group size effect. Our interpretation is that group size (as well as MPCR effects) operate by affecting the weight of beliefs in the decision to contribute which is consistent with previous literature (e.g., Weimann et al. (2019); Kerr (1989)).

9 Summary

In this paper we revisited a classic question in the experimental economics literature on public goods games: the role of group size and MPCR for voluntary cooperation. Our starting point was the observation that many people have a preference for conditional cooperation (Fischbacher et al., 2001; Fischbacher and Gächter, 2010; Chaudhuri, 2011; Thöni and Volk, 2018), which renders the beliefs about others' cooperation important. The specific question we addressed in this paper was how group size and MPCR affect preferences (that is, attitudes to cooperation) and beliefs, and how preferences and beliefs jointly explain cooperation.

For our analysis we used the ABC approach which measures an individual i's attitudes to cooperation (a_i) and beliefs (b_i) to explain i's cooperation (c_i): $a_i(b_i) \rightarrow c_i$, which has proven to be a good framework to explain, quantitatively, the level of cooperation we see in small-group experimental public goods games. The ABC approach we used in this paper $[a_i(f, b_i(f)) \rightarrow c_i]$ allows to go beyond simply observing resulting levels of cooperation as a function of the contextual features of the public good game: the contextual features f (group size and MPCR parameters in our case) might affect a_i or b_i or both, and it is their interactions $a_i(b_i)$ that explains the cooperation levels c_i we observe.

The ABC approach requires measuring the attitudes to cooperation a_i, for which (Fischbacher et al., 2001) (FGF) introduced an incentive-compatible method. The ABC approach has been successfully used in small groups of four players with very similar MPCR parameters (Fischbacher and Gächter, 2010; Fischbacher et al., 2012;

Gächter et al., 2017, 2022; Isler et al., 2021). Until this paper, there has been no evidence on how the ABC approach fares in larger groups and with different MPCR parameters. Moreover, the FGF method to measure attitudes to cooperation might not be appropriate in large groups because it loses its incentive-compatibility as group size increases. We therefore also introduced a new incentive-compatible version of the strategy method, which is scalable to any group size. Thus, in addition to applying the ABC approach to investigate MPCR and group size effects, the Scalable Strategy Method is our second contribution in this paper.

We conducted two studies. In Study 1 we applied the Scalable Strategy Method (SSM) and compared it with the traditional FGF method with group sizes 3 and 9 and MPCRs of 0.4 and 0.8. We found that the ABC approach on average correctly predicted actual contribution levels in all conditions. In terms of predictive success and using cross-validated mean squared error and R^2 as criteria, the SSM method did at least as well as FGF, in particular in large groups.

In Study 2, we only used SSM and compared group sizes of 3 and 30, and MPCRs of 0.04, 0.08, 0.4 and 0.8. Again, the ABC approach predicts contribution levels in all conditions.

Our results are striking: none of the game parameters in both studies, nor the elicitation method of cooperative attitudes (FGF or SSM) affects attitudes to cooperation (a_i) in a significant way, with the exception of groups of 30 and low MPCR, where we find a_i to be slightly flatter than in other conditions. In terms of cooperation behavior, we find that higher MPCRs lead to higher beliefs and higher cooperation, confirming previous evidence. Holding MPCR constant, we find no significant evidence for group size effects: group size neither affects a_i, nor b_i, nor c_i. However, group size magnifies the effects of beliefs on cooperation. Similarly, we conclude that MPCR effects work via increased beliefs about others' cooperativeness and not via changed attitudes to cooperation.

In summary, we conclude that the ABC approach, coupled with the scalable strategy method, works well to explain cooperation for a wide range of parameter sets. Future research should therefore explore the applicability of the ABC approach and the scalable strategy method in further public goods settings of economic interest.

Acknowledgements

This work was supported by the European Research Council [Grant Number ERC-AdG 101020453 PRINCIPLES] and funded by the Centre of Decision Research and Experimental Economics (CeDEx) at the University of Nottingham. We are grateful for comments by Patrick Maus and Kieran Stockley, as well as those received at seminars at the University of Nottingham, the 8th Annual Workshop on Behavioral Game Theory at the University of East Anglia and the International Conference on Social Dilemmas (2022).

Data Availability

Data and analysis files will be posted on https://osf.io/7psud/ upon acceptance.

Online Appendix

The Online Appendix contains the experimental instructions and additional figures.

References

Andreoni, J., "Cooperation in public-goods experiments: kindness or confusion?," *American Economic Review*, 1995a, pp. 891–904.

_ , "Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments," *The Quarterly Journal of Economics*, 1995b, *110* (1), 1–21.

_ **, M. A. Kuhn, and C. Sprenger**, "Measuring time preferences: A comparison of experimental methods," *Journal of Economic Behavior & Organization*, 2015, *116*, 451–464.

Arechar, A., S. Gächter, and L. Molleman, "Conducting interactive experiments online," *Experimental Economics*, 2018, *21* (1), 99–131.

Bayer, R., E. Renner, and R. Sausgruber, "Confusion and learning in the voluntary contributions game," *Experimental Economics*, 2013, *16* (4), 478–496.

Brandts, J. and A. Schram, "Cooperation and noise in public goods experiments: applying the contribution function approach," *Journal of Public Economics*, 2001, *79* (2), 399–427.

_ **and G. Charness**, "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, 2011, *14* (3), 375–398.

Burton-Chellew, M., C. El Mouden, and S. West, "Conditional cooperation and confusion in public-goods experiments," *Proceedings of the National Academy of Sciences*, 2016, *113* (5), 1291–1296.

Carpenter, J. P., "Punishing free-riders: How group size affects mutual monitoring and the provision of public goods," *Games and Economic Behavior*, 2007, *60* (1), 31–51.

Chang, L., A. Smith, M. Dufwenberg, and A. Sanfey, "Triangulating the neural, psychological, and economic bases of guilt aversion," *Neuron*, 2011, *70* (3), 560–572.

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

Croson, R., "Theories of commitment, altruism and reciprocity: Evidence from linear public goods games," *Economic Inquiry*, 2007, *45* (2), 199–216.

Diederich, J., T. Goeschl, and I. Waichman, "Group size and the (in) efficiency of pure public good provision," *European Economic Review*, 2016, *85*, 272–287.

Drouvelis, M., *Social preferences: an introduction to behavioural economics and experimental research*, Agenda Publishing, 2021.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt, "The framing of games and the psychology of play," *Games and Economic Behavior*, 2011, *73* (2), 459–478.

Eichenseer, M. and J. Moser, "Conditional cooperation: Type stability across games," *Economics Letters*, 2020, *188*, 108941.

Fehr, E. and I. Schurtenberger, "Normative foundations of human cooperation," *Nature Human Behaviour*, 2018, *2* (7), 458–468.

_ **and K. Schmidt**, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

Fischbacher, U. and S. Gächter, "Social preferences, beliefs, and the dynamics of free riding in public goods experiments," *American Economic Review*, 2010, *100* (1), 541–56.

_ **,** _ **, and E. Fehr**, "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 2001, *71* (3), 397–404.

_ **,** _ **, and S. Quercia**, "The behavioral validity of the strategy method in public good experiments," *Journal of Economic Psychology*, 2012, *33* (4), 897–913.

Fosgaard, T. R., L. G. Hansen, and E. Wengström, "Framing and misperception in public good experiments," *The Scandinavian Journal of Economics*, 2017, *119* (2), 435–456.

Gächter, S. and B. Herrmann, "Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2009, *364* (1518), 791–806.

_ and _ , "The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia," *European Economic Review*, 2011, *55* (2), 193–210.

_ , F. Kölle, and S. Quercia, "Reciprocity and the tragedies of maintaining and providing the commons," *Nature Human Behaviour*, 2017, *1* (9), 650–656.

_ , F. Kölle, and S. Quercia, "Preferences and perceptions in Provision and Maintenance public goods," *Games and Economic Behavior*, 2022, *135*, 338–355.

Goeree, J. K., C. A. Holt, and S. K. Laury, "Private costs and public benefits: unraveling the effects of altruism and noisy behavior," *Journal of Public Economics*, 2002, *83* (2), 255–276.

Guttman, J., "Matching behavior and collective action: Some experimental evidence," *Journal of Economic Behavior & Organization*, 1986, *7* (2), 171–198.

Hergueux, J. and N. Jacquemet, "Social preferences in the online laboratory: a randomized experiment," *Experimental Economics*, 2015, *18* (2), 251–283.

Houser, D. and R. Kurzban, "Revisiting kindness and confusion in public goods experiments," *American Economic Review*, 2002, *92* (4), 1062–1069.

Isaac, R. and J. Walker, "Group size effects in public goods provision: The voluntary contributions mechanism," *The Quarterly Journal of Economics*, 1988, *103* (1), 179–199.

Isaac, R. M., J. M. Walker, and A. W. Williams, "Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups," *Journal of Public Economics*, 1994, *54* (1), 1–36.

Isler, O., S. Gächter, A. Maule, and C. Starmer, "Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics," *Scientific Reports*, 2021, *11* (1), 1–17.

Katuščák, P. and T. Miklánek, "What drives conditional cooperation in public good games?," *Experimental Economics*, 2022, pp. 1–33.

Kerr, N., "Illusions of efficacy: The effects of group size on perceived efficacy in social dilemmas," *Journal of Experimental Social Psychology*, 1989, *25* (4), 287–313.

Keser, C. and H. Kliemt, "The Strategy Method as an instrument for the exploration of limited rationality in oligopoly game behavior (Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes)(by Reinhard Selten)," in "The Art of Experimental Economics," Routledge, 2021, pp. 30–36.

Ledyard, J. O., "Public goods: A survey of experimental research," *Handbook of Experimental Economics*, 1995, *883* (111-194), 884.

List, J. A., "Young, selfish and male: Field evidence of social preferences," *The Economic Journal*, 2004, *114* (492), 121–149.

Marwell, G. and R. E. Ames, "Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem," *American Journal of Sociology*, 1979, *84* (6), 1335–1360.

Mill, W. and M. Theelen, "Social value orientation and group size uncertainty in public good dilemmas," *Journal of Behavioral and Experimental Economics*, 2019, *81*, 19–38.

Mullett, T., R. L. McDonald, and G. Brown, "Cooperation in Public Goods Games Predicts Behavior in Incentive-Matched Binary Dilemmas: Evidence for Stable Prosociality," *Economic Inquiry*, 2020, *58* (1), 67–85.

Nosenzo, D., S. Quercia, and M. Sefton, "Cooperation in small groups: the effect of group size," *Experimental Economics*, 2015, *18* (1), 4–14.

Palan, S. and C. Schitter, "Prolific. ac—A subject pool for online experiments," *Journal of Behavioral and Experimental Finance*, 2018, *17*, 22–27.

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

Pereda, M., I. Tamarit, A. Antonioni, J. Cuesta, P. Hernández, and A. Sánchez, "Large scale and information effects on cooperation in public good games," *Scientific Reports*, 2019a, *9* (1), 1–10.

__ , V. Capraro, and A. Sánchez, "Group size effects and critical mass in public goods games," *Scientific Reports*, 2019b, *9* (1), 1–10.

Schonlau, M., "CROSSVALIDATE: Stata module to perform k-fold crossvalidation," Statistical Software Components, Boston College Department of Economics 2020.

Selten, R., "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," *Beiträge zur Experimentellen Wirtschaftsforschung, Tübingen*, 1967, pp. 136–168.

Spadaro, G., I. Tiddi, S. Columbus, S. Jin, A. Ten Teije, CoDa Team, and D. Balliet, "The Cooperation Databank: Machine-Readable Science Accelerates Research Synthesis," *Perspectives on Psychological Science*, 2022, *17*, 1472–1489.

Sugden, R., "Reciprocity: the supply of public goods through voluntary contributions," *The Economic Journal*, 1984, *94* (376), 772–787.

Thöni, C. and S. Volk, "Conditional cooperation: Review and refinement," *Economics Letters*, 2018, *171*, 37–40.

Volk, S., C. Thöni, and W. Ruigrok, "Temporal stability and psychological foundations of cooperation preferences," *Journal of Economic Behavior & Organization*, 2012, *81* (2), 664–676.

Weber, T. O., B. Beranek, F. Lambarraa, S. Gächter, and J. F. Schulz, "The Behavioral Mechanisms of Voluntary Cooperation in WEIRD and Non-WEIRD societies," *CeDEx Discussion Paper*, 2021, (2021-03).

Weber, T., O. Weisel, and S. Gächter, "Dispositional free riders do not free ride on punishment," *Nature Communications*, 2018, *9* (1), 1–9.

Weimann, J., J. Brosig-Koch, T. Heinrich, H. Hennig-Schmidt, and C. Keser, "Public good provision by large groups–the logic of collective action revisited," *European Economic Review*, 2019, *118*, 348–363.

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

Appendix

Here we describe some further technical details and illustrate with an example how the Scalable Strategy Method works. The instructions are available in the Online Appendix and a Qualtrics version (Qualtrics qsf format) is available at https://osf.io/7psud/.

Procedure: Scalable Strategy Method (SSM)

Consider a group of three people playing a PGG. Table A.1 displays sample responses to the unconditional contribution and the contribution table (i.e. first 11 rows). For illustrative purposes, let's say we have a conditional cooperator, a free rider and a triangle cooperator. Our randomizing device can result in each player playing with the unconditional contribution or with the contribution table. The steps are the following:

- First: Calculate the *computed conditional contribution* for i by plugging in the average contribution of the remaining players into player i's table. We can think of the contribution table as a function that depends on the average contributions of the other players. Example: Player 1 will have a computed conditional contribution of 2 because the average of the other two players unconditional contributions is $(0 + 4)/2 = 2$. Then if we plug in 2 in Player 1's table we get a computed conditional contribution of 2. With the same reasoning Player 2 and Player 3's computed conditional contribution will be 0 and 1, respectively.

- Second: Randomize for each player i to use either their unconditional contribution or their conditional table (with 50% probability).

  - If the unconditional contribution is chosen, we use all other player's computed conditional contributions to determine i's individual payoff. Example: If Player 1 is assigned to the unconditional contribution, he/she would contribute 5. To compute player i's payoff we also need the contributions of the other players. For those we will use the computed conditional contributions. Example of Player 1's payoff: $10 - 5 + \frac{\alpha(5+0+2)}{3}$.
  - If the conditional table is chosen, we use player i's computed conditional contribution and all other player's unconditional contributions to determine i's individual payoff. Example: Player 1 contributes 2. To compute player i's payoff we also need the contributions of the other players. For those, we will use the unconditional contributions. Example of Player 1's payoff: $10 - 2 + \frac{\alpha(2+0+4)}{3}$.

– If for all players their unconditional contribution is selected, we calculate the public goods payoffs from the unconditional contribution. If the conditional contribution is selected for all players, payoffs are calculated from the computed conditional contributions.

Type examples	Player 1 Conditional Cooperator	Player 2 Free Rider	Player 3 Triangle Cooperator
Unconditional cont.	5	0	4
Conditional table			
0	0	0	0
1	1	0	1
2	2	0	1
3	3	0	2
4	4	0	3
5	4	0	4
6	5	0	3
7	5	0	3
8	6	0	2
9	6	0	2
10	7	0	1
Computed conditional cont.	(2)	(0)	(2)

Table A.1: Example of responses. The parenthesis in the last row indicate that the numbers are computed by the researcher using the other rows as explained in the procedure.