Journal of Economic Behavior and Organization 223 (2024) 1–20

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research paper

Testing isomorphic invariance across social dilemma games

Irene Maria Buso [a], Lorenzo Ferrari [b,1], Werner Güth [c], Luisa Lorè [d,*], Lorenzo Spadoni [e]

[a] Department of Economics, University of Bologna, Italy
[b] Italian Competition Authority, Italy
[c] Max Planck Institute for Research on Collective Goods, Germany
[d] Department of Economics, Universität Innsbruck, Austria
[e] Department of Economics and Law, University of Cassino and Southern Lazio, Italy

ARTICLE INFO

JEL classification:
C71
C92
D70
D90

Keywords:
Social dilemma experiments
Isomorphic invariance
Public goods game
Prisoners' dilemma game
Voluntary cooperation

ABSTRACT

Can purely behavioral aspects affect voluntary cooperativeness in isomorphic social dilemma games? We experimentally test isomorphic invariance by comparing frames whose identical payoffs are described as of the Prisoners' Dilemma or the linear Public Good. Participants play two consecutive rounds of the same frame, with no between-round feedback information, interacting with either the same or a different number of other subjects in each round. Hence, frames are compared between subjects whereas games with different numerosity are compared both within subjects and between subjects. Our analysis rejects isomorphic invariance and shows significantly lower average cooperativeness in the linear Public Good frame across all conditions. Moreover, we find a significantly negative effect of group size, especially in Prisoners' Dilemma.

1. Introduction

Isomorphic invariance is an independence requirement that forbids arbitrariness in determining perfectly rational decision-making and game-playing. Yet, several contributions in the experimental literature provide evidence that strategically irrelevant aspects such as, for instance, the naming of the game and of the actions available, and how incentives are presented to players, can trigger behavioral changes in social dilemma games. These effects are found both in Prisoner's Dilemma (e.g., Liberman et al., 2004; Zhong et al., 2007; Ellingsen et al., 2012; Pruitt, 1967, 1970) and in Public Good games (e.g., Elliott et al., 1998; Dufwenberg et al., 2011; Khadjavi and Lange, 2015; Cartwright and Ramalingam, 2019; Andreoni, 1995; Park, 2000).

In particular, although experimenters might be willing to induce a normal form game in which participants are informed exclusively about the choice sets and material payoffs of all players, they can never be sure which game is implemented experimentally (e.g., Cason and Plott, 2014). While participants' comprehension of actual incentives can effectively be enhanced by having them play trial rounds and controlled for by asking control questions, the dilemma-specific presentation of the experimental task can affect the psychological perception of the game even when subjects correctly understand the game form. In particular, it has been shown that the context in which a social dilemma is embedded can make more prominent different psychological motives

* Corresponding author.
E-mail addresses: irenemaria.buso@unibo.it (I.M. Buso), lorenzo.ferrari@agcm.it (L. Ferrari), gueth@coll.mpg.de (W. Güth), luisa.lore@uibk.ac.at (L. Lorè), lorenzo.spadoni@unicas.it (L. Spadoni).

[1] The views and opinions expressed in this article are strictly those of the author and they do not reflect in any way those of the Institution to which he is affiliated.

https://doi.org/10.1016/j.jebo.2024.04.024
Received 25 May 2023; Received in revised form 20 February 2024; Accepted 23 April 2024
Available online 16 May 2024

as care and anger or kindness, affect the level of conflict of interest perceived in the game and foster individualistic reasoning rather than "we-reasoning" (Butler et al., 2011; Butler, 2012; Columbus et al., 2020; Gächter et al., 2022; Ring et al., 2023).

It has been shown that presenting the task in different ways may produce differences in social preferences, beliefs on others' behavior, perception of incentives as gains and losses (the so-called *value framing*), and normative interpretations of the game (see Gerlach and Jaeger, 2016, who review possible determinants of framing effects). These perceptions shape beliefs and preferences determining cooperativeness. Experimentally one may limit such influences via "clean" or neutral framing. Nevertheless, neutral framing has been shown to cause yet another framing effect (Engel and Rand, 2014; Eriksson and Strimling, 2014). Overall, this kind of experimental evidence calls into question the behavioral validity of isomorphic invariance not only when considering different implementations of the same social dilemma game, but also across social dilemma games.

While differences in cooperativeness between isomorphically invariant Common Pool Resources and Public Good games have been extensively investigated in the literature, little attention has been devoted so far to the behavioral comparison of other isomorphically invariant social dilemmas (see Cartwright and Ramalingam, 2019 for a recent review). In this study we fill this gap by testing the behavioral validity of isomorphic invariance across two social dilemmas, i.e., the Prisoners' Dilemma (hereafter, PD) and linear the Public Good (hereafter, PG). Both PD and PG are workhorses for studying cooperation in social science, and a great amount of evidence has been provided using these two settings. Yet, the extent to which one can generalize the findings on cooperativeness and treatment effects (for example, the effect of group size) across PD and PG is questioned by the systematically different implementations of these two games in terms of incentive structures, design features and task presentations (Goetze, 1994). As our experimental setting guarantees that the two frames have the very same incentive structure, we are able to test for the existence of frame-dependent behavioral effects on cooperativeness. Our experimental analysis rejects the hypothesis of isomorphic invariance between PD and PG. In particular, we find that subjects contribute significantly less, and are more likely to completely free ride, in PG. Interestingly, this game frame negatively affects the players' expectations concerning the choice of their co-players. In other words, not only do subjects cooperate less, but also think that the other group members will do so in PG. Additionally, we test the robustness of the treatment effect when changing group size and degree of familiarity with the setting of play. In both game frames subjects play two rounds without feedback with either the same number of players (2 or 3) or with a different number of players in each round. We find that the effect on cooperation of the game frame emerges more clearly in the smaller group as increasing the group size has a greater negative impact on cooperation in PD than in PG. Also, greater familiarity with the setting tends to increase cooperativeness in PG more than it does in PD, thereby reducing the game frame effect.

Section 2 provides states our main research questions and a detailed description of the experimental games; results are presented in Section 3, followed by conclusions in Section 4.

2. Experimental design and research questions

PD and PG games provide paradigmatic examples of social dilemmas and have been widely investigated, both theoretically and empirically, in disciplines that study cooperativeness. However, these two settings, which can be shown to be isomorphically invariant, are typically embedded with payoff-relevant and irrelevant dilemma-specific characteristics. Goetze (1994) has provided a compelling review of the most common features of the experimental implementation of PD and PG, thus allowing us to identify their main peculiarities. First, the strategies available to players in the two games are typically named differently, i.e., "defect" in PD and "not contributing" in PG.[2] Moreover, the nature of the non-excludable good produced via cooperation and the collective action needed to produce it are normally explicitly mentioned in PG but not in PD. The way incentives are presented is also different, as payoffs in PD and PG are generally expressed in terms of final wealth and changes in wealth, respectively. Finally, the two games are typically characterized by a different number of available choices and players, and this in turn affects their incentive structures. In particular, multi-player, multiple-choice PG, and two-player, binary-choice PD have been more typically implemented in the experimental literature.

Given the above-mentioned differences between PD and PG, we now introduce and discuss the main research questions that our isomorphically invariant setting allows us to answer.

Research Question 1 (R.1)

The multiplicity of game-specific features raises caution when comparing findings collected in implementations of PD and PG as these are characterized by systematically different incentive structures and provide different contexts to players, potentially impacting their perception of the game. Implementing two isomorphically equivalent PD and PG games we aim primarily at testing whether the mere difference in task description is able to affect cooperation levels as well as its distribution, i.e., the proportion of free riders. As discussed above, in fact, the description of the experimental task might affect the perception of the two games and trigger different levels of cooperativeness.

While the payoff matrix presented in each game frame is identical due to isomorphic invariance, what crucially distinguishes PD and PG in our experimental test is the difference in the verbal description of the task. Table 1 presents an extract from the instructions for the condition $n = 2$ that exemplifies the difference between the two game frames.

[2] Although payoff equivalent, "defect" may suggest failures to take some positive action, while "not contributing" denotes inaction leading to differences in behavior. Our instructions limit the influence of jargon like "defect" and "not contributing" as we refer to contribute in PG, but we do not contrast it with the term "defection" in PD.

Table 1
Framing of round-specific instructions $n = 2$.

PD	PG
In this round, you will interact with one other participant. You and the other participant can choose among four options: two options are denoted as option o and option a. The other two options, available to each of you, are the intermediate option M_1, on the basis of which 1/3 of option o and 2/3 of option a are used, and the intermediate option M_2, on the basis of which 1/3 of option a and 2/3 of option o are used. [...] For example, if both of you choose option o, respectively a, you will obtain 240, respectively 432. You obtain 456 when you choose o and the other a, and 216 when you choose a and the other o. If you both choose M_1, respectively M_2, each of you obtains 304, respectively 368.	In this round, you will interact with one other participant. Each of you is endowed with 240 tokens. You and the other participant have four choice options specifying how many of these tokens each of you invest in a joint project from which both you both gain equally (i.e., how much each of you gains from the joint project depends only on the sum of both investments in the joint project). [...] For example, if both of you choose to invest 0, respectively 240 tokens, each of you obtains 240, respectively 432. You obtain 456 when you invest 0 tokens and the other 240, and 216 when you invest 240 tokens and the other 0. If you both choose to invest 80, respectively 160 tokens, each of you obtains 304, respectively 368.

These task descriptions are designed to be representative of the main differences, surveyed in the first part of this section, typically encountered in how PD and PG are presented to experimental subjects. These differences could impact the players' psychological perception of the two games, potentially leading to different levels of cooperativeness. Given the several (and in some cases, opposing) effects that the frame could produce, however, it is not possible to hypothesize *ex-ante* whether cooperativeness will be larger in PD or in PG. On the one hand, PG is typically characterized by an explicit reference to the non-excludable and non-subtractable good produced by cooperation, i.e., the "common project", while in PD the production of a non-excludable good is only implicit and, as a consequence, subjects receive no information about its non-subtractable or subtractable nature. Additionally, PG explicitly appeals to a collective action. These two differences might induce a "we-thinking" reasoning in PG (but not in PD), i.e., a reasoning oriented towards the maximization of group rather than individual payoff, and this could lead to greater cooperativeness in PG than in PD (Butler, 2012). Furthermore, since cooperation has been frequently shown to be lower when the public good is subtractable than when it is non-subtractable, the explicit reference to this second case in PG could also sustain greater cooperativeness in this setting rather than in PD (e.g., Gächter et al., 2022, 2017).

On the other hand, the different way incentives are presented in the two games could lead to lower cooperativeness in PG than in PD. In fact, while PD matrices present exclusively the players' payoffs resulting from cooperative or non-cooperative actions, PG payoffs are presented as resulting from subtracting individual contributions from one's endowment and adding what all group members gain from collective action. Referring to initial token endowments in PG may suggest being entitled to it and the subject might then feel less compelled to contribute (see Hoffman and Spitzer, 1985).

In the same vein, defining the payoffs in terms of wealth changes in PG could induce players to think in terms of gains and losses. Contribution could then be discouraged if players have the perception of suffering a loss when subtracting tokens from their own endowment when confronted with the choice to contribute to the public good.

Research Question 2 (R.2)

In some of our sessions, we vary the group size (together with the number of available choices) from $n = 2$ in the first round to $n = 3$ in the second or vice-versa, while in others we keep it constant in both rounds. This allows us to check for the presence of a group size effect and, on top of that, whether this effect is frame-dependent. Although the evidence is mixed for both PG and PD, there is a slightly greater tendency in the literature to find a positive effect in PG and a negative one in PD.[3] Komorita and Lapworth (1982), for instance, provide evidence of a negative effect of group size in a one-shot PD with n varying from 2 to 3. On the other hand, Goeree et al. (2002) shows a tendency toward a positive effect of group size in a one-shot PG with n varying from 2 to 4. Barcelo and Capraro (2015) find a positive impact of group size in a one-shot PG and a negative one in a one-shot (not isomorphically invariant) PD.[4] While the design differences (such as the number of choice options, the *MPCR*, and number of players) in the two settings might explain these puzzling findings, one cannot rule out a possible frame-dependent effect of the number of players.

Our experimental investigation contributes, via guaranteeing isomorphic invariance, to the strand of literature on the effect of group size in social dilemmas and sheds light on whether differences between behavior in PD and PG are likely to depend on different design features or different task descriptions.

Research Question 3 (R.3)

Finally, our experimental design allows us to check whether the round of play, which is a measure of players' familiarity with the task confronted in each experimental frame, affects cooperativeness. In other words, as each subject plays the same game type twice, even if they do so without between-round feedback information, and potentially with a different number of other players in the second round, we can check whether playing twice could affect the cooperativeness level in the second round, i.e., when the subject is more familiar with the experimental task. Since in our experimental design group size and available choices vary in some

3 Nosenzo et al. (2015) and Barcelo and Capraro (2015) offers comprehensive reviews of group size effect in PD and PG.

4 Nosenzo et al. (2015) point out that the positive effect reported in Goeree et al. (2002) does not achieve statistical significance if only treatments with the same external and internal returns are considered.

sessions but not in others, we are able to check whether this familiarity effect depends on the frame (PD or PG), on the number of players in the two rounds, and on whether subjects interact with the same number of other players (2 or 3) in both rounds or with a different number of other players in each round.

We now illustrate in detail the incentive structure of the two games and the experimental design we use to test these hypotheses in detail.

2.1. Isomorphic invariance and experimental design

Our experimental setup guarantees isomorphic invariance, in that the payoff matrices provided in both the PD- and PG-games are identical for each n-condition. The payoff matrix for each condition is presented to the subjects in the round instructions and remains available to them while making their choices. The following equation provides a simple representation of the theoretical framework for our experiment which illustrates the isomorphic invariance between the PD and PG frames. Assume that player i's payoff, denoted with u_i, is defined as

$$u_i = 1 - c_i + \alpha C.$$

The individual choice c_i is the fraction of the unit endowment ($E = 1$) that player i chooses to contribute in PG, while in PD c_i is the chosen share of the full Cooperativeness strategy.[5] Similarly, $C = \sum_{i=1}^{n} c_i$ is the total contribution in PG and the number of cooperators (overall share of E strategies) in PD. As previously mentioned, while the experimental implementations of the PG usually allow for intermediate levels of contribution, PD implementations have been mostly experimentally implemented as a dichotomous choice of cooperating or defecting. Still, experimental implementations of Prisoner's dilemmas as a non-binary choice game have been performed (for instance, Goerg and Walkowitz, 2010 and Goerg et al., 2020 implement it as a continuous choice game). In our experiment, we allow for intermediate levels of contribution in PG as well as to choose the degree of cooperativeness in PD. In addition to the full Cooperativeness ($c_i = 1$) and defection ($c_i = 0$) strategies, players in PD can choose $\frac{1}{3}$ and $\frac{2}{3}$ when $n = 2$ and $\frac{1}{2}$ when $n = 3$ as intermediate shares of full cooperativeness.[6] The payoffs in the matrices are determined according to the following parameters: the initial endowment of points for contributing in PG is 240, which corresponds to the points earned in PD if everyone defects. The other payoffs are calculated according to the formula provided in this section for a value of α equal to 0.9 when $n = 2$ and to 0.6 when $n = 3$.[7] The payoff matrices for each n−condition are presented in Appendix B.

While the payoff matrix shown to players in each game frame is identical due to isomorphic invariance, the task is presented differently in PD and PG, as already illustrated in Table 1.

Altogether the decision data distinguish:

$n \in \{2, 3\}$, the number of players;
$j \in \{PD, PG\}$, the frame; and
$t \in \{1st, 2nd\}$, the round of play.

In all sessions subjects played two one-shot rounds, i.e., subjects received feedback on choices and payoffs only after the last round. The different frames, PD and PG, vary between subjects, while the number of players changed both within subjects and between subjects: in some sessions, subjects played with the same group size – $n = 2$ or $n = 3$ – in both rounds, while in other sessions we let group size vary across rounds. Table 2 illustrates the 8 experimental conditions implemented: 2 (PD, PG) × 2 ($n = 2$, 3 in round 1) × 2 (same n in both rounds, different n in both rounds).

Table 2
Experimental conditions.

	Round 1	Round 2	
PD	$n = 2$	$n = 2$	(same n)
	$n = 3$	$n = 3$	
	$n = 2$	$n = 3$	(different n)
	$n = 3$	$n = 2$	
PG	$n = 2$	$n = 2$	(same n)
	$n = 3$	$n = 3$	
	$n = 2$	$n = 3$	(different n)
	$n = 3$	$n = 2$	

[5] Notice that $c_i = E = 1$ represents full cooperativeness in PD and maximal contribution in PG.

[6] Note that the two conditions may differ in the level of complexity, as $n = 2$ results in a 4 × 4 payoff table, whereas $n = 3$ the latter is 3 × 3 × 3.

[7] We control for efficiency incentives by varying the *MPCR*, $alpha$, such that $\alpha_{n=3} = \frac{2}{3}\alpha_{n=2}$. Note that in the 2- and 3- person versions participants face different numbers of available choices: $n = 2$ results in a 4 × 4 payoff table, whereas $n = 3$ the latter is 3 × 3 × 3; $n = 3$ combines the larger number of players with fewer available options in order to reduce complexity differences across n-conditions.

2.2. Experimental procedure

General instructions are presented before round 1, while specific instructions for each task are shown only before the corresponding round. The general instructions inform subjects that they will play more than one round (without specifying the exact number of rounds) that they will never interact with the same subject(s),[8] and that one round will be randomly drawn for payment. The general instructions also specify the formula for transforming points P, earned in that round, in the probabilities of earning either 4 or 14 euros, so that each participant in the experiment earns either 4 or 14 euros in addition to the show-up fee of 6 euros. Specific instructions inform subjects about the number of players and the game, including the corresponding payoff matrix, to be played in that round. Before actually making a decision in a given round, subjects were asked to answer two control questions (answers could be submitted only after two minutes) and could proceed only after having answered both correctly. If three minutes had elapsed without a correct answer, subjects were provided with feedback about the correct answer and allowed to proceed. In each round subjects selected one of their available choices and stated their non-incentivized beliefs about the choice behavior of the other group member(s). After the experiment participants completed a non-incentivized questionnaire eliciting demographics, self-reported risk attitude, personality traits as well as a six-item cognitive reflection test (Dohmen et al., 2011; Primi et al., 2016; Rammstedt and John, 2007). We conducted the experiment at the Luiss Cesare Lab, with a total of 384 participants in 16 sessions. The first 8 sessions took place in February 2022 and involved 192 participants. In these sessions, the number of players changed between the two rounds, with 48 subjects in each of the four conditions. The second set of 8 sessions was conducted in September 2023 and involved an additional 192 participants. These sessions focused on conditions where the number of players remained constant between the two rounds, with 48 subjects in each of the four conditions.[9] Sessions lasted, on average, 45 min. The average payment of 16.02 euros includes the participation fee of 6 euros. The English version of the Instructions can be found in Appendix B. The experiment was programmed in oTree and involved only students of Luiss University, recruited via ORSEE (Chen et al., 2016; Greiner, 2015).

3. Results

3.1. Main results

We first provide the results for our main research question (R.1), i.e., whether the game frame, PD or PG, affects cooperativeness. We then focus specifically on testing the hypotheses about the effect on cooperativeness of the number of players (R.2) and of the round of play (R.3) and on checking whether the effects are frame-dependent. In the following analyses, we define Cooperativeness as normalized individual cooperativeness in the two frames, with zero and one corresponding respectively to dominant free riding and to the fully cooperative choice.

Table 3 reports average Cooperativeness across game frames, group sizes, and rounds, with a focus on the history of the game in round 2, i.e., whether group size was constant and equal to $n = 2$ or $n = 3$ in both rounds, or changed from $n = 2$ ($n = 3$) to $n = 3$ ($n = 2$) in the second round. Overall, Cooperativeness is significantly higher in PD than in PG (p-value < 0.0001), as also shown in Fig. 1. Focusing on round 1, this difference is driven by groups composed of two players ($n = 3$ is not statistically significant). Furthermore, looking at round 2, results show that average cooperativeness in PD is significantly larger than in PG in all histories of the game except for $n = 3$ in round 2 and $n = 2$ in round 1, where the difference is positive but not significant. Hence, isomorphic invariance can be overall rejected. The seemingly weaker effect on Cooperativeness of the frame when $n = 3$ will be discussed in the analysis of group size, where we show that this has a negative effect on the cooperativeness in both frames, but especially in PD.[10]

Result 1. game frames have a significant effect on average Cooperativeness, which is significantly higher in PD than in PG. Hence, isomorphic invariance between PD and PG is rejected

We now assess the effect of group size on average Cooperativeness. As second-round choices might in principle be affected by the different histories of the game, we focus on the first round. Table 3 shows that Cooperativeness is consistently larger in $n = 2$ than in $n = 3$ in both frames. This result is confirmed by a t-test where the round-1 sub-sample is split by group size (not shown in the table, p-value $= 0.0002$). This result is mainly driven by PD, for which the difference is very large and highly significant (0.59 in $n = 2$ versus 0.38 in $n = 3$, p-value $= 0.0005$).[11]

Result 2. The number of players has a significantly negative effect on average Cooperativeness. This effect is stronger in PD than in PG

[8] Player groups of 2 and 3 are formed from matching groups of six participants to determine the final payment.

[9] Two anonymous referees raised valid concerns about the original sample size and order effects, prompting us to conduct the second set of additional sessions with a constant group size. These sessions not only allowed us to drastically increase statistical power but also to address order effects that had previously confounded our main results.

[10] Figs. 3, and 4 in Appendix A provide a graphical representation of Table 3.

[11] The difference is smaller and only mildly significant in PG (0.42 versus 0.32, and p-value $= 0.0727$).

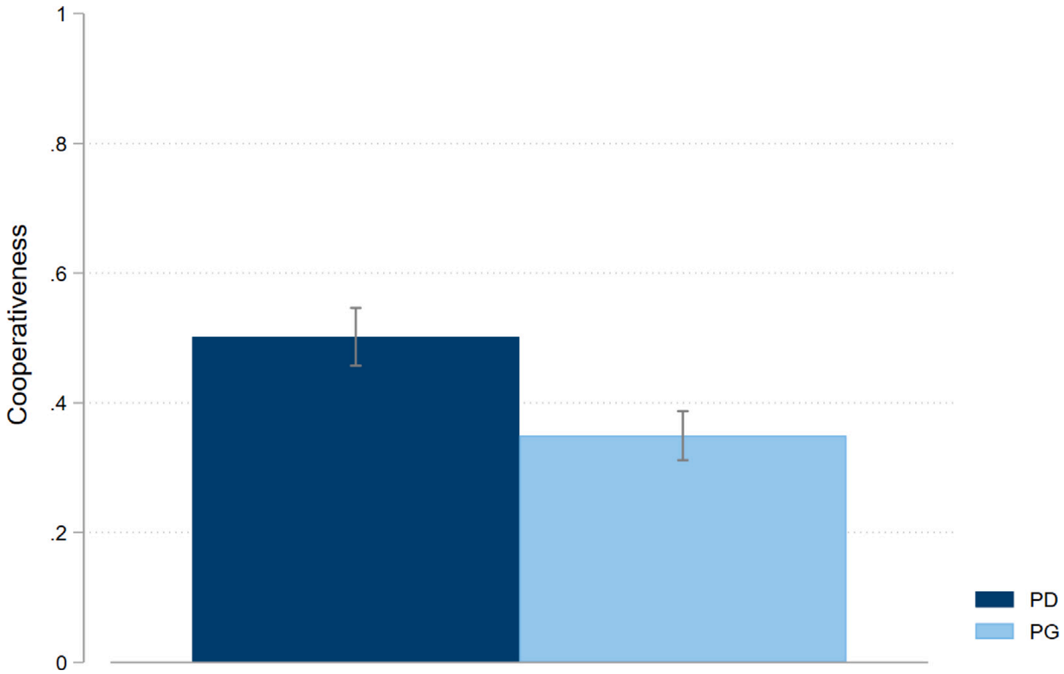

Fig. 1. Average cooperativeness across game frames.
Notes: Confidence intervals on means are at 95%.

Table 3
Frame effect on Cooperativeness.

		mean		p-value
		PD	PG	
	all rounds[a]	0.502 (0.444) [384]	0.349 (0.378) [384]	**< 0.0001**
round 1	$n = 2$	0.594 (0.428) [96]	0.424 (0.382) [96]	**0.0041**
	$n = 3$	0.375 (0.429) [96]	0.323 (0.391) [96]	0.38
round 2	$n = 2$ (both rounds)	0.66 (0.421) [48]	0.41 (0.359) [48]	**0.0023**
	$n = 2$ ($n = 3$ in round 1)	0.458 (0.427) [48]	0.299 (0.365) [48]	**0.0520**
	$n = 3$ (both rounds)	0.604 (0.449) [48]	0.344 (0.388) [48]	**0.0031**
	$n = 3$ ($n = 2$ in round 1)	0.354 (0.437) [48]	0.25 (0.342) [48]	0.1967

Notes: p-values refer to t tests of the hypothesis that average cooperativeness does not differ in the two groups. Standard deviations are in parentheses. The number of observations is in brackets.

[a] These tests encompass two choices per subject.

Finally, we investigate the effect on the Cooperativeness of the round of play. To this purpose, we limit our focus to subjects who have played with the same number of other participants in both rounds. Changing group size between the two rounds, in fact, also entails varying the number of choices available to players, thus impairing comparison. The difference in Cooperativeness between rounds 1 and 2 for this sub-sample of subjects is only mildly significant (0.43 in round 1 versus 0.5 in round 2, and p-value = 0.1005). When testing separately the two game frames, moreover, we fail to reject the null hypothesis in both PD (p-value = 0.1416) and PG (p-value = 0.3682).

Result 3. Repeating the game a second time has a positive and *mildly significant*, not frame-dependent effect on average Cooperativeness

3.2. Econometric evidence

Our main findings show that Cooperativeness is significantly affected by the game frame, the group size, and the round of play. Including these variables in a regression analysis seems a natural way to shed additional light on some aspects that are potentially not captured by hypothesis testing.

Results are reported in Table 4. In line with the previous section, column (1) confirms the significantly negative effect of PG and $n = 3$ on Cooperativeness, while round of play displays a negligible and not significant coefficient. Moreover, controlling for the interactions between the main variables, as shown in column (2), allows us to better capture the effect of familiarity on Cooperativeness: Subjects appear to cooperate significantly less in round 2 if they face a different n-condition than in round 1, but significantly more if the n-condition does not change. This additional piece of evidence allows us to provide a better interpretation for Result 3. Even without receiving between-round feedback information, subjects familiar with the choice design in round 2 (same group size and available choice set as in round 1) tend to be more cooperative, whereas players facing a different choice design in round 2 tend to be less cooperative. In columns (3)-(6) of Table 4 we further disentangle the effects on Cooperativeness by frame. columns 3 and 4 show that, in PD, $n = 3$ has a significantly negative effect on Cooperativeness. Moreover, we find a mildly significant and positive effect of experiencing the same n-condition. No significant effect is found for the second round. As shown in columns 5 and 6, the round of play is statistically significant in PG, both alone and in its interaction with the same n, with opposite signs as described earlier in the overall regression. $n = 3$ is still negative but with a coefficient is less significant than in PD.[12] These additional findings thus allow us to enrich our main results and to confirm the overall significantly negative effect on the Cooperativeness of the PG-frame. An increase in group size, from 2 to 3, has a negative effect, and this effect is especially significant in PD and mildly so in PG. The effect of the round of play depends on the game history. Round 2 has a mildly significant negative effect on Cooperativeness, while its interaction with same n is positive and strongly significant. Hence, as previously stated in Result 3, repeating the game while keeping the group size constant has a positive effect on Cooperativeness. However, this further result warns us about the possibility of having the opposite effect on Cooperativeness when the task is repeated with a different group size. This last result suggests that Cooperativeness is decreased by frame-specific familiarity and increased by game-specific familiarity. This effect is driven by the PG-frame.

Table 4
Main determinants of Cooperativeness.

	All frames		PD		PG	
	(1)	(2)	(3)	(4)	(5)	(6)
PG	−0.152***	−0.170***				
	(0.032)	(0.059)				
$n = 3$	−0.114***	−0.219***	−0.149***	−0.219***	−0.079**	−0.101*
	(0.029)	(0.062)	(0.044)	(0.062)	(0.038)	(0.056)
Round 2	−0.007	−0.112*	0.035	−0.094	−0.048	−0.166***
	(0.027)	(0.066)	(0.042)	(0.072)	(0.034)	(0.060)
PG × $n = 3$		0.118				
		(0.083)				
PG × Round 2		−0.035				
		(0.080)				
$n = 3$ × Round 2		0.139		0.139		0.043
		(0.089)		(0.089)		(0.078)
PG × $n = 3$ × Round 2		−0.095				
		(0.119)				
Same n		0.009		0.108*		−0.090
		(0.042)		(0.062)		(0.056)
Same n × Round 2		0.155***		0.118		0.193***
		(0.053)		(0.082)		(0.066)
Observations	768	768	384	384	384	384

Notes: Dependent variable. *Cooperativeness*: variable ranging between 0 and 1. Regressors. *PG*: 1 if PG and 0 if PD. *Round 2*: 1 if round 2 and 0 otherwise; *n=3*: 1 if group size is 3 and 0 if 2. Specifications are OLS regressions with clustered standard errors at subject level reported in parentheses.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

To disentangle how the game frame affects the distribution of Cooperativeness and, in particular, free riding behavior (i.e., $c_i = 0$), in Table 5 we present two Hurdle regressions, separately for sessions with constant and varying n-conditions. We chose to run separate regression as we are interested in analyzing the effect of $n = 3$ on both the extensive and intensive margin, which is not correctly estimated when $n = 3$ changes across rounds.[13] In both models, the dependent variables in columns 1 and 2 are respectively a dummy variable equal to 1 if $c_i > 0$ for $i = 1, 2$ (i.e., $y = 1$ if $c_1 > 0 \& c_2 > 0$) and the level of contribution conditional on the fact that this is positive (i.e., $y = c_i | c_i \geq 0$). In other words, columns (1) and (3) show the effects of round-invariant covariates on

[12] To account for other variables that might have an effect on Cooperativeness, we have included a richer specification of the three models presented in this section in Table 7 in Appendix A. Self-reported risk tolerance and being male have a mildly significant positive effect on overall Cooperativeness respectively in PG and PD, while the coefficients included in the main specification are not qualitatively affected.

[13] In particular, the Hurdle model does not allow for the inclusion of round-changing variables as controls in the first hurdle, as the first hurdle has just one outcome per subject, thus making the two types of sessions not comparable.

the probability of a positive choice, i.e., the probability of abstaining from complete free riding (contributing zero in both rounds), while columns (2) and (4) the effect of the same controls on cooperativeness, provided that this is positive.

Our findings show that, when the n-condition does not change, PG makes subjects significantly more likely not only to cooperate less (as already discussed), but even to free ride. The effect of PG on complete free riding, however, is not significant with varying n-condition. Furthermore, $n = 3$ positively affects the likelihood of free riding exclusively when n stays constant. Hence, it appears that the overall negative effect found on $n = 3$ is driven by greater free riding in sessions where the n-condition does not change.

Result 4. PG not only has a significantly negative effect on Cooperativeness but also on its distribution in sessions with same n. Moreover, the negative effect of group size on Cooperativeness seems driven by the greater free riding

Table 5
The effect of framing on the distribution of Cooperativeness.

	same n		different n	
	First Hurdle (1)	Second Hurdle (2)	First Hurdle (3)	Second Hurdle (4)
PG	−.232***	−0.222***	−0.740	−0.115***
	(0.044)	(0.034)	(0.045)	(0.044)
$n = 3$	−0.138***	−0.018		0.068*
	(0.043)	(0.050)		(0.036)
Round 2		−0.011		−0.029
		(0.044)		(0.037)
$n = 3$ × Round 2		0.073		
		(0.067)		
Observations	384		384	

Notes: Regressions (1–3) and (2–4) are the first and second hurdle of a double hurdle model with clustered standard errors at subject level in the second equations. The dependent variables in (1–3) and (2–4) are respectively a dummy variable equal to 1 when cooperativeness is positive, and the level of cooperativeness when cooperativeness is strictly positive. Regressors. *PG*: 1 if PG and 0 if PD. *n=3*: 1 if group size is 3 and 0 if 2. *Round 2*: 1 if round 2 and 0 otherwise.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

3.3. A focus on choices motivation: Beliefs

We finally analyze the effect of players' beliefs about their co-player's (if $n = 2$) or two co-players' (if $n = 3$) choice on Cooperativeness. As expected, average beliefs[14] are overall highly correlated with Cooperativeness (correlation= 0.4551). To account for possible opportunistic behavior, we denote *Delta* the difference between average beliefs and individual Cooperativeness. Fig. 2 illustrates the average *Delta* levels across game frame, round, and group size. Unsurprisingly, the differences are always positive: self-serving subjects cooperate, on average, less than what they expect the others will. Furthermore, delta is significantly larger in $n = 3$ overall (0.19 versus 0.11, p-value = 0.0079), in PG (0.18 versus 0.10, p-value = 0.0324), but only mildly in PD (0.20 versus 0.12, p-value = 0.0977).

Columns (1) and (2) in Table 6 investigate the effects of the game frame on average beliefs regarding co-players' Cooperativeness. In line with our previous results on choice, we find that players expect others to cooperate less in PG than in PD. Moreover, when including our full set of covariates, average beliefs are significantly affected by $n = 3$ both alone (negatively) and when interacting with the other controls. In particular, it appears that the positive coefficients of the interactions between $n = 3$ and PG, and between $n = 3$ and round, are mainly driven by the significant effects, respectively negative and positive, of $n = 3$ and of the same interaction term in PD, as shown in column (5).[15,16]

Result 5. Subjects expect their co-players to cooperate significantly less in PG than in PD and when $n = 3$. The effect of group size is mainly driven by PD

[14] Average beliefs correspond to a player's beliefs about the choice of the only other co-players when $n = 2$ and to the average of a player's beliefs about the choice of the other two co-players when $n = 3$.

[15] To account for other variables that might have an effect on average beliefs, we have included a richer specification of the three models presented in this section in Table 8 in Appendix A. Personal characteristics seem to play a role in determining the subjects' expectations about co-players' Cooperativeness while not qualitatively affecting the results found in the main specification. Interestingly, self-reported risk tolerance (at least weakly) significantly increases beliefs while a higher score in a CRT test decreases them in all specifications. Moreover, age has a significantly negative effect in all specifications. This could be indicative of sophistication resulting from greater experience in participating in experiments or a better understanding of theoretical predictions, particularly if subjects have prior exposure to courses involving game theory. Finally, people who self-reported as more extroverted and agreeable (conscientious) tend to display significantly higher (lower) beliefs.

[16] Appendix A displays the frequencies of choices concerning cooperativeness in relation to subjects' beliefs about the behavior of the other co-player(s).

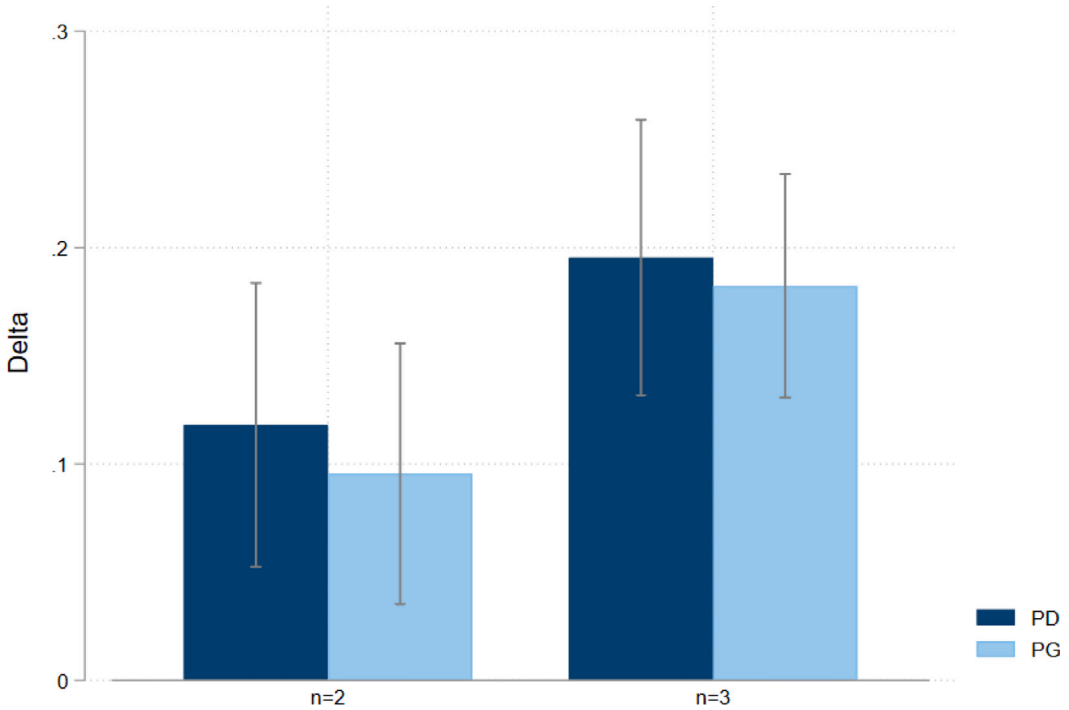

Fig. 2. Delta across game frame, rounds, and group size.
Notes: *Delta* is the difference between average beliefs and individual Cooperativeness.
Confidence intervals on means are at 95%.

Table 6
Main determinants of Beliefs.

	All frames		PD		PG	
	(1)	(2)	(3)	(4)	(5)	(6)
PG	−0.170***	−0.247***				
	(0.031)	(0.058)				
$n = 3$	−0.032	−0.194***	−0.072*	−0.194***	0.008	0.071
	(0.028)	(0.057)	(0.040)	(0.057)	(0.039)	(0.054)
Round 2	0.031	−0.071	0.087**	−0.076	−0.025	0.006
	(0.023)	(0.060)	(0.034)	(0.065)	(0.032)	(0.063)
PG × $n = 3$		0.265***				
		(0.078)				
PG × Round 2		0.073				
		(0.075)				
$n = 3$ × Round 2		0.243***		0.243***		−0.127*
		(0.077)		(0.077)		(0.074)
PG × $n = 3$ × Round 2		−0.370***				
		(0.106)				
Same n		−0.024		0.006		−0.054
		(0.039)		(0.057)		(0.054)
Round 2 × Same n		0.073		0.082		0.064
		(0.047)		(0.068)		(0.064)
Observations	768	768	384	384	384	384

Notes: Dependent variable. *Average Belief*: variable ranging between 0 and 1. Regressors. *PG*: 1 if PG and 0 if PD. *n=3*: 1 if group size is 3 and 0 if 2. *Round 2*: 1 if round 2 and 0 otherwise. *Same n*: 1 if same number of players in both rounds and 0 otherwise. Specifications are OLS regressions with clustered standard errors at subject level reported in parentheses.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4. Conclusions

In our experimental analysis, we aim to assess the validity of isomorphic invariance in social dilemma games. Our between-subjects design involves a comparison of the Prisoner's Dilemma and linear Public Good game frames with identical payoffs. We also exogenously manipulate the group size, allowing subjects to participate in two rounds within the same frame, but without receiving any feedback in between. This setup enables us to explore not only the impact of the game frame on cooperativeness but also the influence of group size and familiarity.

We provide evidence against isomorphic invariance. On average, subjects exhibit significantly less cooperation and are more inclined to free-ride in the Public Good game when compared to the Prisoner's Dilemma. Additionally, the condition with a larger group size significantly reduces cooperativeness, while players tend to cooperate more when confronted with the same choice structure and group size in the second round. This familiarity effect is predominantly observed among those playing the Public Good game. Furthermore, isomorphic invariance does not hold behaviorally even when looking at players' expectations regarding cooperativeness choices of the others.

One plausible explanation for the observed decrease in cooperativeness within the Public Good frame can be linked to the way incentives are presented. In contrast to the Prisoner's Dilemma, the Public Good game grants participants an initial token endowment. The distribution of property rights implied by this endowment may discourage cooperativeness, as subjects might feel entitled to the ownership of these tokens (Hoffman and Spitzer, 1985). Existing literature has also suggested that the initial allocation of tokens to a *Private* account can lead to different contribution behaviors compared to allocation to a *Group* account, where deductions can be made, or to an *Investment* account, where tokens must be distributed between a *Private* and a *Group* account. Khadjavi and Lange (2015) shows that the initial allocation of tokens significantly influences the proportion of free-riders. In this regard, Cartwright (2016) has highlighted that the initial allocation of tokens is a design element often manipulated to establish a "give-take" frame in public good games, often overlooking the potential impact this manipulation may have on cooperation. Therefore, our primary finding concerning the game frame would suggest that the initial assignment of tokens to a *Private* account, a characteristic of the Public Good game but not the Prisoner's Dilemma,[17] has a negative impact on cooperativeness.

Another aspect of incentive presentation that may contribute to the negative effect of the Public Good frame is the definition of payoffs in terms of wealth changes in Public Good, as opposed to final wealth in Prisoner's Dilemma. Expressing wealth changes in relation to an endowment in Public Good might encourage individuals to consider the endowment their reference point and think in terms of gains and losses. Consequently, losses from contributing may appear more relevant than gains from cooperation, in line with the principles of prospect theory (Kahneman and Tversky, 1979).

Additionally, our experimental analysis makes a significant contribution to the literature on group size effects in social dilemmas. We have demonstrated a negative effect when transitioning from 2 to 3 players in both game frames. This effect primarily results from the impact of group size on the choice to free-ride, rather than on the level of cooperativeness. This negative effect aligns with previous evidence indicating that when the setting already strongly encourages cooperation (high *MPCR*), increasing the group size is more likely to have a negative impact (Nosenzo et al., 2015). Moreover, our study has revealed that the negative effect of group size is more pronounced in Prisoner's Dilemma than in Public Good. This observation may explain why previous research has shown a negative group size effect in Prisoner's Dilemma (e.g., Komorita and Lapworth, 1982; Barcelo and Capraro, 2015), while a positive effect has been shown to predominate in Public Good Games.[18] Since our findings indicate that familiarity significantly influences Cooperativeness, particularly within the Public Good frame, this should serve as a cautionary note for researchers who interpret effects across rounds solely based on the feedback they provide. Such interpretations might be confounded by the influence of familiarity, which can operate even in the absence of feedback. Based on our evidence, such concerns should not pose a significant issue when dealing with Prisoner's Dilemma designs. In addition, our study has revealed that task repetition yields different outcomes in terms of cooperativeness depending on whether the number of players changes or remains constant. Given that task repetition is a common practice in one-shot settings to improve subjects' understanding of incentives, this finding underscores the need to reevaluate the effectiveness of this practice and the sensitivity of subjects' comprehension. In particular, considering the potential role that misperceptions of incentives might play in generating framing effects in social dilemmas (e.g., Gächter et al., 2022), we believe that future research should delve deeper into the potential misperceptions in both Prisoner's Dilemma and Public Good scenarios and assess the effectiveness of increasing game understanding through repetition.

Replication files

The data and code for replicating the results of this paper are available at https://osf.io/juysc/. All files are licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Werner Güth would like to acknowledge the Max Planck Institute for Research on Collective Goods, and Luisa Lorè would like to acknowledge "Austrian Science Fund, SFB F63", for supporting this study.

[17] The initial assignment of tokens to a *Private* account "can be recognized as the workhorse of public good experiments" ((Cartwright, 2016) pag. 80).

[18] Note that, according to the considerations derived by Nosenzo et al. (2015), the weaker negative effect that we find in Public Good Game could have turned to a positive effect if we would have had a lower MPCR.

Appendix A. Additional tables and figures

A.1. Cooperativeness

See Figs. 3 and 4 and Table 7.

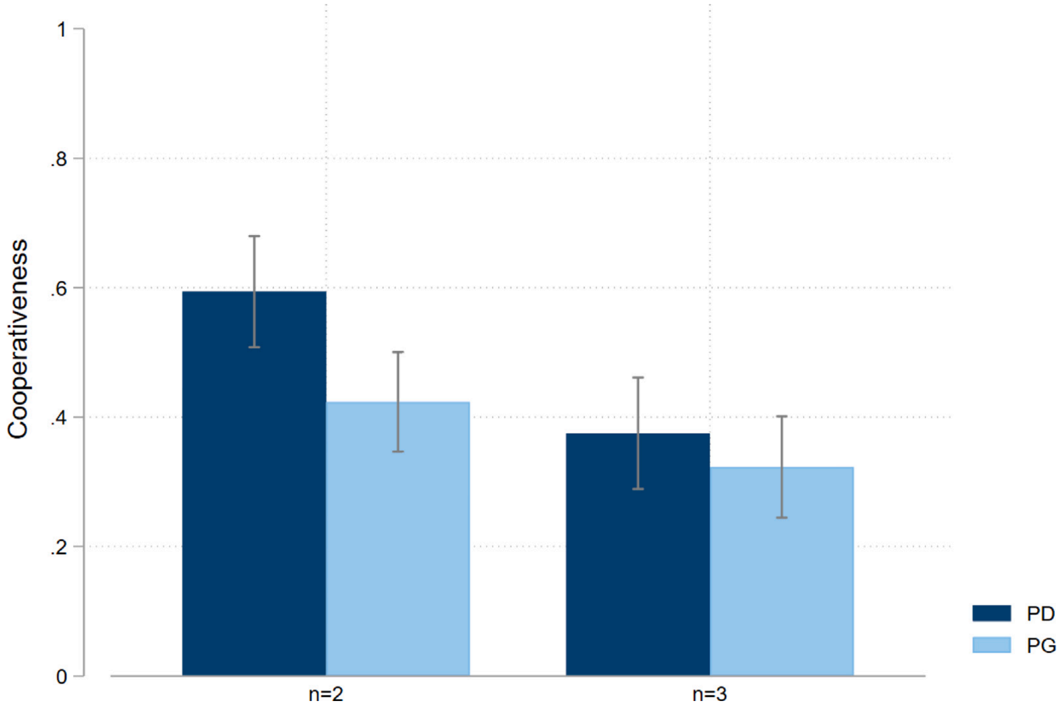

Fig. 3. Average cooperativeness in round 1 across game frames, and player numbers.
Notes: Confidence intervals on means are at 95%.

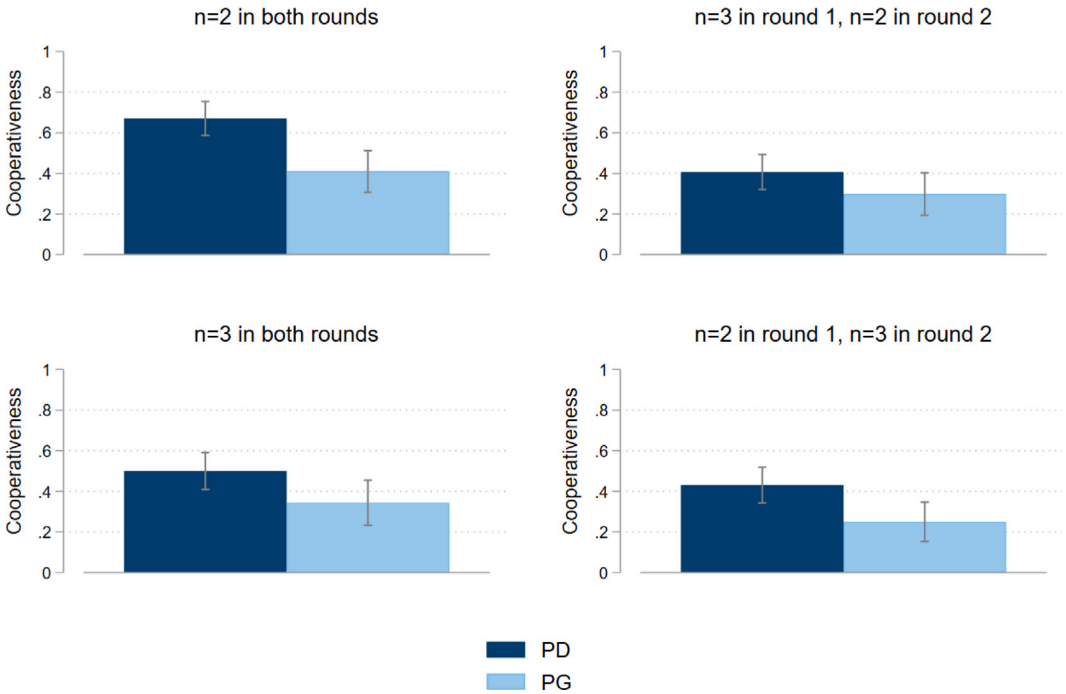

Fig. 4. Average cooperativeness in round 2 across game frames, player numbers, and game history.
Notes: Confidence intervals on means are at 95%.

Table 7
Overall determinants of Cooperativeness.

	All frames (1)	PD (2)	PG (3)
PG	−0.163***		
	(0.059)		
$n = 3$	−0.209***	−0.199***	−0.097*
	(0.061)	(0.061)	(0.055)
Round 2	−0.107	−0.084	−0.163***
	(0.067)	(0.074)	(0.060)
PG × $n = 3$	0.114		
	(0.083)		
PG × Round 2	−0.036		
	(0.080)		
$n = 3$ × Round 2	0.128	0.119	0.037
	(0.090)	(0.091)	(0.079)
PG × $n = 3$ × Round 2	−0.092		
	(0.120)		
Same n	0.001	0.113*	−0.096*
	(0.043)	(0.065)	(0.057)
Round 2 × Same n	0.155***	0.118	0.193***
	(0.053)	(0.084)	(0.067)
Male	0.040	0.109**	−0.031
	(0.036)	(0.055)	(0.049)
Age	−0.003	0.006	−0.013
	(0.008)	(0.012)	(0.011)
Economics	−0.040	−0.033	−0.060
	(0.034)	(0.049)	(0.047)
Risk lover	0.070*	0.035	0.091*
	(0.036)	(0.053)	(0.048)
CRT	−0.015	−0.017	−0.008
	(0.009)	(0.014)	(0.013)
Extraversion	−0.002	0.003	0.003
	(0.009)	(0.013)	(0.012)
Conscientiousness	0.001	0.014	−0.015
	(0.010)	(0.014)	(0.014)
Agreeableness	0.002	−0.013	0.018
	(0.010)	(0.013)	(0.014)
Neuroticism	0.010	0.018	0.002
	(0.008)	(0.011)	(0.010)
Openness	−0.001	0.001	−0.004
	(0.009)	(0.014)	(0.012)
Observations	768	384	384

Notes: Dependent variable. *Cooperativeness*: variable ranging between 0 and 1. Regressors. *PG*: 1 if PG and 0 if PD. *n=3*: 1 if group size is 3 and 0 if 2. *Round 2*: 1 if round 2 and 0 otherwise. *Same n*: 1 if same number of players in both rounds and 0 otherwise. *Individual Controls*: male, age, dummy for studying economics, dummy for self-reported risk attitude, cognitive reflection test (CRT), and personality traits from the Big Five. Specifications are OLS regressions with clustered standard errors at subject level reported in parentheses.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.2. Beliefs

See Fig. 5 and Table 8.

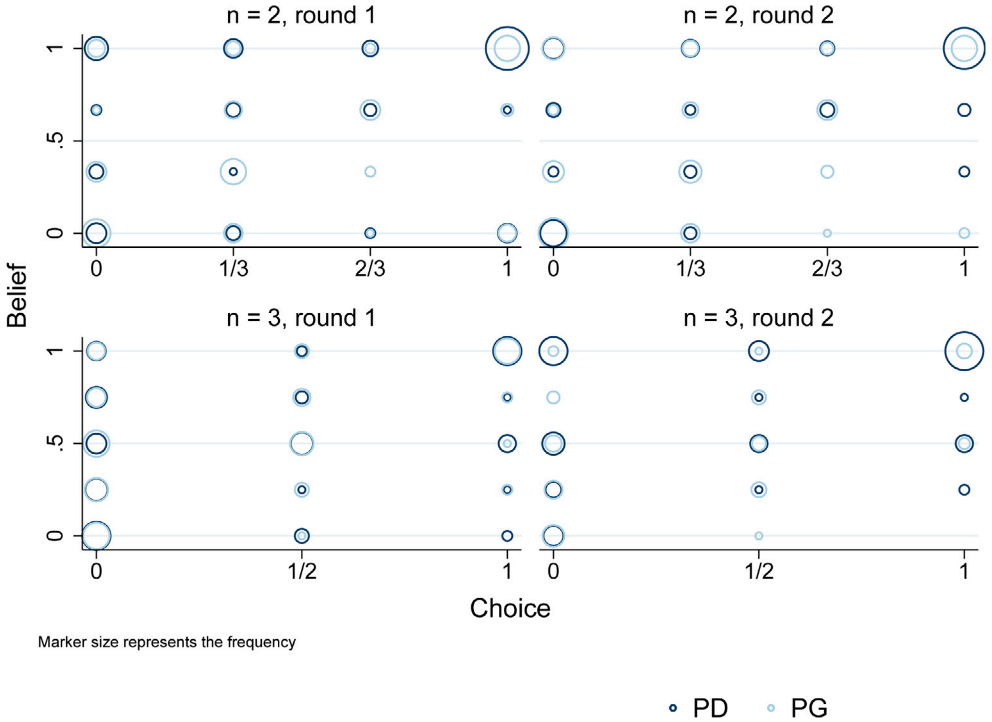

Fig. 5. Frequency of Belief in Others' Cooperativeness by Own Cooperativeness Level.

Table 8
Overall determinants of Beliefs.

	All frames (1)	PD (2)	PG (3)
PG	−0.233***		
	(0.058)		
$n = 3$	−0.190***	−0.178***	0.069
	(0.057)	(0.057)	(0.052)
Round 2	−0.067	−0.064	0.004
	(0.060)	(0.065)	(0.061)
PG × $n = 3$	0.267***		
	(0.078)		
PG × Round 2	0.070		
	(0.074)		
$n = 3$ × Round 2	0.234***	0.220***	−0.123*
	(0.076)	(0.077)	(0.069)
PG × $n = 3$ × Round 2	−0.364***		
	(0.104)		
Same n	−0.021	0.014	−0.049
	(0.040)	(0.060)	(0.053)
Round 2 × Same n	0.073	0.082	0.064
	(0.047)	(0.069)	(0.065)
Male	0.007	0.094*	−0.093**
	(0.034)	(0.050)	(0.045)
Age	−0.012	−0.004	−0.029***
	(0.008)	(0.011)	(0.011)
Economics	−0.0003	0.012	−0.025
	(0.033)	(0.048)	(0.046)
Risk lover	0.123***	0.100*	0.137***
	(0.037)	(0.059)	(0.047)
CRT	−0.020**	−0.019	−0.018
	(0.009)	(0.014)	(0.013)
Extraversion	0.008	0.003	0.023*
	(0.008)	(0.012)	(0.012)
Conscientiousness	−0.004	0.014	−0.026**
	(0.010)	(0.015)	(0.013)
Agreeableness	0.013	0.002	0.027**
	(0.009)	(0.013)	(0.013)
Neuroticism	0.007	0.018	−0.005
	(0.008)	(0.012)	(0.010)
Openness	0.003	0.0003	0.002
	(0.009)	(0.013)	(0.011)
Observations	768	384	384

Notes: Dependent variable. *(Average) Belief*: variable ranging between 0 and 1. Regressors. *PG*: 1 if PG and 0 if PD. *n=3*: 1 if group size is 3 and 0 if 2. *Round 2*: 1 if round 2 and 0 otherwise. *Same n*: 1 if same number of players in both rounds and 0 otherwise. *Individual Controls*: male, age, dummy for studying economics, dummy for self-reported risk attitude, cognitive reflection test (CRT), and personality traits from the Big Five. Specifications are OLS regressions with clustered standard errors at subject level reported in parentheses.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.3. Balance table

See Table 9.

Table 9
Balance table.

Variable	PD	PG	Pairwise t-test Mean difference
Male	0.474	0.500	−0.026
	(0.036)	(0.036)	
Age	20.896	21.156	−0.260
	(0.148)	(0.137)	
Economics	0.620	0.641	−0.021
	(0.035)	(0.035)	
Risk lover	0.760	0.734	0.026
	(0.031)	(0.032)	
CTR	2.901	3.135	−0.234
	(0.129)	(0.128)	
Difficulty	0.854	0.870	−0.016
	(0.026)	(0.024)	
Time choice 1	92.823	81.615	11.208
	(5.570)	(5.040)	
Time choice 2	73.453	58.281	15.172**
	(4.977)	(3.932)	
Observations	192	192	384

Notes: *Balance table*. The values displayed for t-tests are the differences in the means across the groups. Standard errors in parentheses. Variables. male, age, dummy for studying economics, dummy for self-reported risk attitude, cognitive reflection test (CRT), dummy for self-reported difficulty, seconds to choose cooperativeness level in round 1 and in round 2.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.4. Time analysis

See Table 10.

Table 10
Response times.

	Choice time (1)	Total time (2)
PG	−1.348	−1.827
	(10.007)	(10.045)
$n = 3$	26.704**	24.834**
	(11.242)	(11.210)
Round 2	−3.676	−2.080
	(9.903)	(10.051)
PG × $n = 3$	−19.436	−19.102
	(14.918)	(14.974)
PG × $n = 3$	−14.851	−17.346
	(12.176)	(12.344)
$n = 3$ × Round 2	−25.579*	−25.757*
	(15.152)	(15.251)
PG × $n = 3$ × Round 2	21.774	26.171
	(19.978)	(20.130)
Same n	−4.404	−5.086
	(7.710)	(7.730)
Round 2 × Same n	−5.807	−5.885
	(7.231)	(7.316)
Male	−18.796***	−19.599***
	(6.771)	(6.855)
Age	0.150	0.343
	(1.600)	(1.558)
Risk lover	−2.759	−1.809
	(7.426)	(7.469)
CRT	1.134	1.431
	(1.785)	(1.790)
Economics	6.605	7.353
	(6.496)	(6.526)
Extraversion	−2.374	−2.250
	(1.722)	(1.716)
Conscientiousness	1.486	1.455
	(1.821)	(1.842)
Agreeableness	1.392	1.591
	(1.744)	(1.760)
Neuroticism	−1.484	−1.468
	(1.503)	(1.514)
Openness	0.448	0.534
	(1.659)	(1.659)
Observations	768	768

Notes: Dependent variables. In (1), *seconds to choose cooperativeness level*. In (2), *seconds to choose cooperativeness level and beliefs*. Regressors. *PG*: 1 if PG and 0 if PD. *n=3*: 1 if group size is 3 and 0 if 2. *Round 2*: 1 if round 2 and 0 otherwise. *Same n*: 1 if same number of players in both rounds and 0 otherwise. *Individual Controls*: male, age, dummy for studying economics, dummy for self-reported risk attitude, cognitive reflection test (CRT), and personality traits from the Big Five. Specifications are OLS regressions with clustered standard errors at subject level reported in parentheses.
Significance of coefficients: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix B. Instructions

General instructions

Welcome to this experiment.

In this experiment you can earn either 4 or 14 euro in addition to your show-up fee of 6 euro. Whether you earn 4 or 14 euros depends on P, the payoff that you will gain from the experiment. Whatever your payoff P will be, your probability of earning either 4 or 14 euro will be positive and smaller than 100%. You will win 14 euro with probability

$$(P - 100) \times \frac{100}{500}\%,$$

and 4 euro with the complementary probability

$$100\% - (P - 100) \times \frac{100}{500}\%.$$

Your payoff P in each round depends on your choices and the choices of other participant(s) with whom you are interacting. Only one round will be randomly selected by the computer for you payment. All participants are paid in this way, i.e., all participants earn 4 or 14 euro according to their own payoff P.

In each round, you will interact with randomly chosen participants whose identity we will not reveal to you. In Round 1 you will interact with one other participant ($n = 2$) [two other participants ($n = 3$)]. In round 1, you can choose among four (if n=2) [among three (if $n = 3$)] choice options. You will not interact with the same other participant(s) in more than one round. We now explain how choices determine payoffs P of the interacting participants. You will be informed about your own payoff P of the playoff-relevant round only after the last round.

Prisoners' dilemma, $n = 2$

Round 1: [In this round you will interact with one other participant. You and the other participant can choose among four options:]

Round 2: [In this second round you will interact only with one other participant. Differently from round 1, you and the other participant can now choose among four options:]

Two options are denoted as option o and option a. The other two options, available to each of you, are the intermediate option M_1, on the basis of which 1/3 of option o and 2/3 of option a are used, and the intermediate option M_2, on the basis of which 1/3 of option a and 2/3 of option o are used.

Your and the other player's results, expressed in terms of payoff P, are listed in the following table, for each of the $2 \times 2 \times 4 = 16$ possible combinations of your individual choices.

For example, if both of you choose option o, respectively a, you will obtain 240, respectively 432. You obtain 456 when you choose o and the other a, and 216 when you choose a and the other o. If you both choose M_1, respectively M_2, each of you obtains 304, respectively 368.

At the moment of the decision you may resort to this table, which will be shown on the screen. Moreover, in the same screen, by clicking on the INSTRUCTIONS icon at the top right, you can read again the instructions of the round.

Your choice is	The other's choice is	Your payoff P is	The other's payoff P is
o	o	240	240
	M_1 (1/3 of o and 2/3 of a)	312	232
	M_2 (2/3 of o and 1/3 of a)	384	224
	a	456	216
M_1 (1/3 of o and 2/3 of a)	o	232	312
	M_1 (1/3 of o and 2/3 of a)	304	304
	M_2 (2/3 of o and 1/3 of a)	376	296
	a	448	288
M_2 (2/3 of o and 1/3 of a)	o	224	384
	M_1 (1/3 of o and 2/3 of a)	296	376
	M_2 (2/3 of o and 1/3 of a)	368	368
	a	440	360
a	o	216	456
	M_1 (1/3 of o and 2/3 of a)	288	448
	M_2 (2/3 of o and 1/3 of a)	360	440
	a	432	432

Prisoners' dilemma, $n = 3$

Round 1: [In this round you will interact with two other participants. You and the other participants can choose among three options:]

Round 2: [In this second round you will interact with two other participants. Differently from round 1, you and the other participants can choose among three options:]

two options are denoted as option o and option a. The other option, available to each of you, is the intermediate option M, on the basis of which 1/2 of option o and 1/2 of option a are used.

Your and the other players' results, expressed in terms of payoff P, are listed in the following table, for each of the $3 \times 3 \times 3 = 27$ possible combinations of your individual choices.

For example, if all three of you choose option o, respectively a, you will obtain 240, respectively 432. You obtain 528 when you choose o and both others a, and 144 when you choose a and both others o. You obtain 384 when you and another choose o and the third a and 288 when you and another choose a and the third o. If you all choose M, each of you obtains 336.

At the moment of the decision you may resort to this table, which will be shown on the screen. Moreover, in the same screen, by clicking on the INSTRUCTIONS icon at the top right, you can read again the instructions of the round.

Your choice is	The others' choices are	Your payoff P is	The others' payoff P are
o	Both play o	240	Both earn 240
	One plays o, the other plays M (1/2 of o and 1/2 of a)	312	One earns 312, the other 192
	One plays o, the other plays a	384	One earns 384, the other 144
	Both play M (1/2 of o and 1/2 of a)	384	Both earn 264
	One plays M (1/2 of o and 1/2 of a), the other plays a	456	One earns 336, the other 216
	Both play a	528	Both earn 288
M (1/2 of o and 1/2 of a)	Both play o	192	Both earn 312
	One plays o, the other plays M (1/2 of o and 1/2 of a)	264	One earns 384, the other 264
	One plays o, the other plays a	336	One earns 456, the other 216
	Both play M (1/2 of o and 1/2 of a)	336	Both earn 336
	One plays M (1/2 of o and 1/2 of a), the other plays a	408	One earns 408, the other 288
	Both play a	480	Both earn 360
a	Both play o	144	Both earn 384
	One plays o, the other plays M (1/2 of o and 1/2 of a)	216	One earns 456, the other 336
	One plays o, the other plays a	288	One earns 528, the other 288
	Both play M (1/2 of o and 1/2 of a)	288	Both earn 408
	One plays M (1/2 of o and 1/2 of a), the other plays a	360	One earns 480, the other 360
	Both play a	432	Both earn 432

Public good, $n = 2$

Round 1: [In this round you will interact with one other participant. Each of you is endowed with 240 tokens. You and the other participants have four choice options specifying how many of these tokens each of you invests in a joint project from which both you both gain equally (i.e., how much each of you gains from the joint project depends only on the sum of both investments in the joint project).]

Round 2: [In this second round you will interact only with one other participant. Each of you is endowed with 240 tokens. Differently from round 1, you and the other participants have four choice options specifying how many of these tokens each of you invests in a joint project from which both you both gain equally (i.e., how much each of you gains from the joint project depends only on the sum of both investments in the joint project).]

Your payoff P is given by the sum of tokens that you choose not to invest in the common project plus what each of you obtains from the overall investment in the common project.

You and the other participant have the same four choice options concerning the investment in the joint project: two options are 0 and 240, according to which respectively no or all available tokens are invested. The other options are the intermediate options 80 and 160, according to which respectively 1/3 and 2/3 of the endowment of tokens is invested.

Your and the other player's results, expressed in terms of payoff P, are listed in the following table, for each of the $2 \times 2 \times 4 = 16$ possible combinations of your individual choices.

For example, if both of you choose to invest 0, respectively 240 tokens, each of you obtains 240, respectively 432. You obtain 456 when you invest 0 tokens and the other 240, and 216 when you invest 240 tokens and the other 0. If you both choose to invest 80, respectively 160 tokens, each of you obtains 304, respectively 368.

At the moment of the decision you may resort to this table, which will be shown on the screen. Moreover, in the same screen, by clicking on the INSTRUCTIONS icon at the top right, you can read again the instructions of the round.

Your choice is	The other's choice is	Your payoff P is	The other's payoff P is
0	0	240	240
	80	312	232
	160	384	224
	240	456	216
80	0	232	312
	80	304	304
	160	376	296
	240	448	288
160	0	224	384
	80	296	376
	160	368	368
	240	440	360
240	0	216	456
	80	288	448
	160	360	440
	240	432	432

Public good, $n = 3$

Round 1: [In this round you will interact with two other participants. Each of you is endowed with 240 tokens. You and the other participants have three choice options specifying how many of these tokens each of you invests in a joint project from which both you both gain equally (i.e., how much each of you gains from the joint project depends only on the sum of both investments in the joint project).]

Round 2: [In this second round you will interact with two other participants. Each of you is endowed with 240 tokens. Differently from round 1, you and the other participants have three choice options specifying how many of these tokens each of you invests in a joint project from which both you both gain equally (i.e., how much each of you gains from the joint project depends only on the sum of both investments in the joint project).]

Your payoff P is given by the sum of tokens that you choose not to invest in the common project plus what each of you obtains from the overall investment in the common project.

You and the other participants have the same three choice options concerning the investment in the joint project: two options are 0 and 240, according to which respectively no or all available tokens are invested. The other options is the intermediate options 120, according to which 1/2 of the endowment of tokens is invested.

Your and the other players' results, expressed in terms of payoff P, are listed in the following table, for each of the $3 \times 3 \times 3 = 27$ possible combinations of your individual choices.

For example, if all three of you choose to invest 0, respectively 240 tokens, you will obtain 240, respectively 432. You obtain 528 when you invest 0 tokens and both others 240, and 144 when you invest 240 tokens and both others invest 0. You obtain 384 when you and another choose to invest 0 and the third invests 240, and 288 when you and another choose to invest 240 tokens and the third 0. If you all choose to invest 120 tokens, each of you obtains 336.

At the moment of the decision you may resort to this table, which will be shown on the screen. Moreover, in the same screen, by clicking on the INSTRUCTIONS icon at the top right, you can read again the instructions of the round.

Your choice is	The others' choices are	Your payoff P is	The others' payoff P are
0	Both play 0	240	Both earn 240
	One plays 0, the other plays 120	312	One earns 312, the other 192
	One plays 0, the other plays 240	384	One earns 384, the other 144
	Both play 120	384	Both earn 264
	One plays 120, the other plays 240	456	One earns 336, the other 216
	Both play 240	528	Both earn 288
120	Both play *o*	192	Both earn 312
	One plays 0, the other plays 120	264	One earns 384, the other 264
	One plays 0, the other plays 240	336	One earns 456, the other 216
	Both play 120	336	Both earn 336
	One plays 120, the other plays 240	408	One earns 408, the other 288
	Both play 240	480	Both earn 360
240	Both play 0	144	Both earn 384
	One plays 0, the other plays 120	216	One earns 456, the other 336
	One plays 0, the other plays 240	288	One earns 528, the other 288
	Both play 120	288	Both earn 408
	One plays 120, the other plays 240	360	One earns 480, the other 360
	Both play 240	432	Both earn 432

References

Andreoni, J., 1995. Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. Q. J. Econ. 110 (1), 1–21.
Barcelo, H., Capraro, V., 2015. Group size effect on cooperation in one-shot social dilemmas. Sci. Rep. 5 (1), 7937.
Butler, D.J., 2012. A choice for 'me'or for 'us'? Using we-reasoning to predict cooperation and coordination in games. Theory and Decision 73, 53–76.
Butler, D.J., Burbank, V.K., Chisholm, J.S., 2011. The frames behind the games: Player's perceptions of prisoners dilemma, chicken, dictator, and ultimatum games. J. Socio-Econ. 40 (2), 103–114.
Cartwright, E., 2016. A comment on framing effects in linear public good games. J. Econ. Sci. Assoc. 2 (1), 73–84.
Cartwright, E., Ramalingam, A., 2019. Framing effects in public good games: Choices or externalities? Econom. Lett. 179, 42–45.
Cason, T.N., Plott, C.R., 2014. Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. J. Polit. Econ. 122 (6), 1235–1270.
Chen, D.L., Schonger, M., Wickens, C., 2016. oTree—An open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Finance 9, 88–97.
Columbus, S., Münich, J., Gerpott, F.H., 2020. Playing a different game: Situation perception mediates framing effects on cooperative behaviour. J. Exp. Soc. Psychol. 90, 104006.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. J. Eur. Econom. Assoc. 9 (3), 522–550.
Dufwenberg, M., Gächter, S., Hennig-Schmidt, H., 2011. The framing of games and the psychology of play. Games Econom. Behav. 73 (2), 459–478.
Ellingsen, T., Johannesson, M., Mollerstrom, J., Munkhammar, S., 2012. Social framing effects: Preferences or beliefs? Games Econom. Behav. 76 (1), 117–130.
Elliott, C.S., Hayward, D.M., Canon, S., 1998. Institutional framing: Some experimental evidence. J. Econ. Behav. Organ. 35 (4), 455–464.
Engel, C., Rand, D.G., 2014. What does "clean" really mean? The implicit framing of decontextualized experiments. Econom. Lett. 122 (3), 386–389.
Eriksson, K., Strimling, P., 2014. Spontaneous associations and label framing have similar effects in the public goods game. Judgm. Decis. Mak. 9 (5), 360–372.
Gächter, S., Kölle, F., Quercia, S., 2017. Reciprocity and the tragedies of maintaining and providing the commons. Nat. Hum. Behav. 1 (9), 650–656.
Gächter, S., Kölle, F., Quercia, S., 2022. Preferences and perceptions in Provision and Maintenance public goods. Games Econom. Behav. 135, 338–355.
Gerlach, P., Jaeger, B., 2016. Another frame, another game?: Explaining framing effects in economic games. In: Proceedings of Norms, Actions, Games. NAG 2016.
Goeree, J.K., Holt, C.A., Laury, S.K., 2002. Private costs and public benefits: unraveling the effects of altruism and noisy behavior. J. Public Econ. 83 (2), 255–276.
Goerg, S.J., Rand, D., Walkowitz, G., 2020. Framing effects in the prisoner's dilemma but not in the dictator game. J. Econ. Sci. Assoc. 6 (1), 1–12.
Goerg, S.J., Walkowitz, G., 2010. On the prevalence of framing effects across subject-pools in a two-person cooperation game. J. Econ. Psychol. 31 (6), 849–859.
Goetze, D., 1994. Comparing prisoner's dilemma, commons dilemma, and public goods provision designs in laboratory experiments. J. Confl. Resolut. 38 (1), 56–86.
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1 (1), 114–125.
Hoffman, E., Spitzer, M.L., 1985. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. J. Leg. Stud. 14 (2), 259–297.
Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.
Khadjavi, M., Lange, A., 2015. Doing good or doing harm: experimental evidence on giving and taking in public good games. Exp. Econ. 18 (3), 432–441.
Komorita, S.S., Lapworth, C.W., 1982. Cooperative choice among individuals versus groups in an N-person dilemma situation. J. Personal. Soc. Psychol. 42 (3), 487.
Liberman, V., Samuels, S.M., Ross, L., 2004. The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. Pers. Soc. Psychol. Bull. 30 (9), 1175–1185.
Nosenzo, D., Quercia, S., Sefton, M., 2015. Cooperation in small groups: the effect of group size. Exp. Econ. 18, 4–14.
Park, E.-S., 2000. Warm-glow versus cold-prickle: a further experimental study of framing effects on free-riding. J. Econ. Behav. Organ. 43 (4), 405–421.
Primi, C., Morsanyi, K., Chiesi, F., Donati, M.A., Hamilton, J., 2016. The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). J. Behav. Decis. Mak. 29 (5), 453–469.
Pruitt, D.G., 1967. Reward structure and cooperation: the decomposed Prisoner's Dilemma game. J. Personal. Soc. Psychol. 7 (1p1), 21.
Pruitt, D.G., 1970. Motivational processes in the decomposed Prisoner's Dilemma game. J. Pers. Soc. Psychol. 14 (3), 227.
Rammstedt, B., John, O.P., 2007. Measuring personality in one minute or less: A 10-item short version of the big five inventory in English and German. J. Res. Pers. 41 (1), 203–212.
Ring, P., Schütt, C.A., Snower, D.J., 2023. Care and anger motives in social dilemmas. Theory and Decision 1–36.
Zhong, C.-B., Loewenstein, J., Murnighan, J.K., 2007. Speaking the same language: The cooperative effects of labeling in the prisoner's dilemma. J. Confl. Resolut. 51 (3), 431–456.