Selfish learning is more important than fair-minded conditional cooperation in public-goods games

Maxwell N. Burton-Chellew*[1,2] & Claire Guérin[2]

*Corresponding author: maxwell.burton@unil.ch

[1]Department of Economics, HEC-University of Lausanne, 1015 Lausanne, Switzerland

[2]Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland.

Version date: Monday 22nd November 2021

ABSTRACT

Why does human cooperation often unravel in economic experiments despite a promising start? Previous studies have interpreted the decline as the reaction of disappointed cooperators retaliating in response to lesser cooperators (conditional cooperation). This interpretation has been considered evidence of a uniquely human form of cooperation, motivated by altruistic concerns for fairness and requiring special evolutionary explanations. However, experiments have typically shown individuals information about both their personal payoff and information about the decisions of their groupmates (social information). Showing both confounds explanations based on conditional cooperation with explanations based on individuals learning how to better play the game. Here we experimentally decouple these two forms of information, and thus these two learning processes, in public goods games involving 616 Swiss university participants. We find that payoff information leads to a greater decline, supporting a payoff-based learning hypothesis. In contrast, social information has small or negligible effect, contradicting the conditional cooperation hypothesis. We also find widespread evidence of both confusion and selfish motives, suggesting that human cooperation is maybe not so unique after all.

Key words: altruism; behavioral economics; confusion; reciprocity; social preferences.

ACKNOWLEDGEMENTS

We are grateful to Tiago Fernandes for assistance with translations and conducting the first study; Luis Santos-Pinto and Laurent Lehmann for helpful discussion, and the audience of the Behavioral Economics and Experimental Research group at UNIL for feedback.

Declarations of interest: none

Highlights

- Human cooperation often unravels in experiments despite a promising start
- We tested if this decline was due to people learning about people, or payoffs
- We find solid support for the role of payoffs, less for learning about people
- We also confirm that many participants are selfish but confused about payoffs
- Our results contradict many interpretations of economic experiments

PsycINFO Classification code: 3000 - Social Psychology

JEL Classification code: DO1 – Microeconomic Behavior: Underlying Principles

INTRODUCTION

"*Sound knowledge about the specific motives behind altruistic acts predominantly stems from laboratory experiments*" - (Fehr & Fischbacher, 2003)

Economic experiments using the public-goods game have shown that many people initially pay costs that benefit their group (cooperate) but that this cooperation is fragile and typically declines over time (Burton-Chellew & West, 2021; Chaudhuri, 2011; Jouxtel, 2019; Ledyard, 1995; Neugebauer, Perote, Schmidt, & Loos, 2009; Thielmann, Böhm, Ott, & Hilbig, 2021; Zelmer, 2003). There are two major competing explanations for this decline that are both based on different forms of learning (Andreoni, 1995; Burton-Chellew & West, 2021; Cooper & Stockman, 2002; Houser & Kurzban, 2002). One hypothesis is that cooperation unravels due to disappointment among cooperative individuals as they learn that people are not as cooperative as they anticipated. This 'conditional cooperators' hypothesis posits that people are mostly motivated by an altruistic concern for fairness, and that human cooperation requires a unique evolutionary explanation (Fehr & Fischbacher, 2003; Fehr & Schurtenberger, 2018; Henrich & Muthukrishna, 2021). The competing hypothesis is that cooperation decreases as confused individuals learn from their payoffs that contributing is not as profitable as they thought. This 'confused learners' hypothesis posits that people are mostly motivated by self-interest but are not well adapted to laboratory games (Burton-Chellew, El Mouden, & West, 2016; Burton-Chellew, Nax, & West, 2015; Burton-Chellew & West, 2013; Burton-Chellew & West, 2021).

Despite decades of study there is still much debate and controversy over these competing explanations, that are not fully mutually exclusive (Burton-Chellew & West, 2021; Camerer, 2013; Fehr & Schurtenberger, 2018). One reason for this is because the standard experimental design typically confounds these two explanations, by coupling together information on a focal player's earnings from the game (payoff information) with information on the decisions of their groupmates (social information). If individuals can observe both payoff and social information, it is unclear which information is more psychologically salient and driving behavior. This is a problem because the two forms of information are often highly correlated: more cooperative groupmates lead to higher payoffs and vice versa. Consequently, both the conditional cooperators and the confused

learners hypotheses can predict an increase/decrease in cooperation in response to more/less cooperation by others (M. N. Burton-Chellew, C. El Mouden, & S. A. West, 2017b; Burton-Chellew, et al., 2015; Nax, Burton-Chellew, West, & Young, 2016).

One solution to the problem of confounding is to use modified versions of the public goods game that test for either Conditional Cooperators or Confused Learners (Andreoni, 1995; Andreozzi, Ploner, & Saral, 2020; Angelovski, Di Cagno, Güth, Marazzi, & Panaccione, 2018; Bayer, Renner, & Sausgruber, 2013; Burton-Chellew, et al., 2016; Burton-Chellew & West, 2013; Croson, Fatas, & Neugebauer, 2005; Di Cagno, Galliera, Güth, & Panaccione, 2016; Ferraro & Vossler, 2010; U. Fischbacher & Gachter, 2010; U. Fischbacher, Gachter, & Fehr, 2001; Gunnthorsdottir, Houser, & McCabe, 2007; Houser & Kurzban, 2002; Kocher, Martinsson, Persson, & Wang, 2016; Kuemmerli, Burton-Chellew, Ross-Gillespie, & West, 2010; Kurzban & Houser, 2005; Shapiro, 2009; Thoni & Volk, 2018). However, while these experiments are informative, it could be argued these modified public goods games are problematic because they are either not the same game or not the same decision mechanism, and thus behavior may change in unknown ways (Camerer, 2013; Urs Fischbacher, Gächter, & Quercia, 2012).

We decoupled payoff information from social information in the standard public goods game and with standard instructions (Figure 1). Our primary method was to provide individuals with either just their personal payoff information (Payoff treatment), or just social information on the decisions of their groupmates (Social treatment, individual decisions were shown in either full, or just the group average decision). This decoupling allowed us to evaluate the relative importance of payoff-based learning and conditional cooperation for causing the decline in contributions. It also had the advantage that we did not have to change the underlying game or basic decision process. If the decline is only due to payoff-based learning, then the decline will not occur unless payoff information is shown i.e., when participants are only shown social information. In contrast, if the decline is only due to conditional cooperation, then adding payoff information will not increase the rate of decline.

As the decline in contributions over time may occur for other reasons, such as increased time for reflection, we included a control treatment to measure how behavior may change over time (Figure 1). This provided a baseline rate of decline for comparison when individuals have neither payoff nor social information (No-info treatment). We also included a fourth treatment that presented the usual, combined, information, replicating the typical design in the literature (Combined treatment). This combined treatment allowed us to evaluate the effect of adding payoff information to social information, which is technically redundant if individuals fully understand the game (a key assumption of the conditional cooperation interpretation). This redundancy is a strength of our design, because if individuals perfectly understand the game and are fair-minded conditional cooperators then behavior should not differ between the two treatments. If, on the other hand, individuals are confused learners, their behavior should differ across treatments, as payoff-based learning will only be possible when payoffs are shown. We also included a range of alternative decision tasks both before and after the main game to test for confusion and motivations more directly (Supplementary Figure 1).

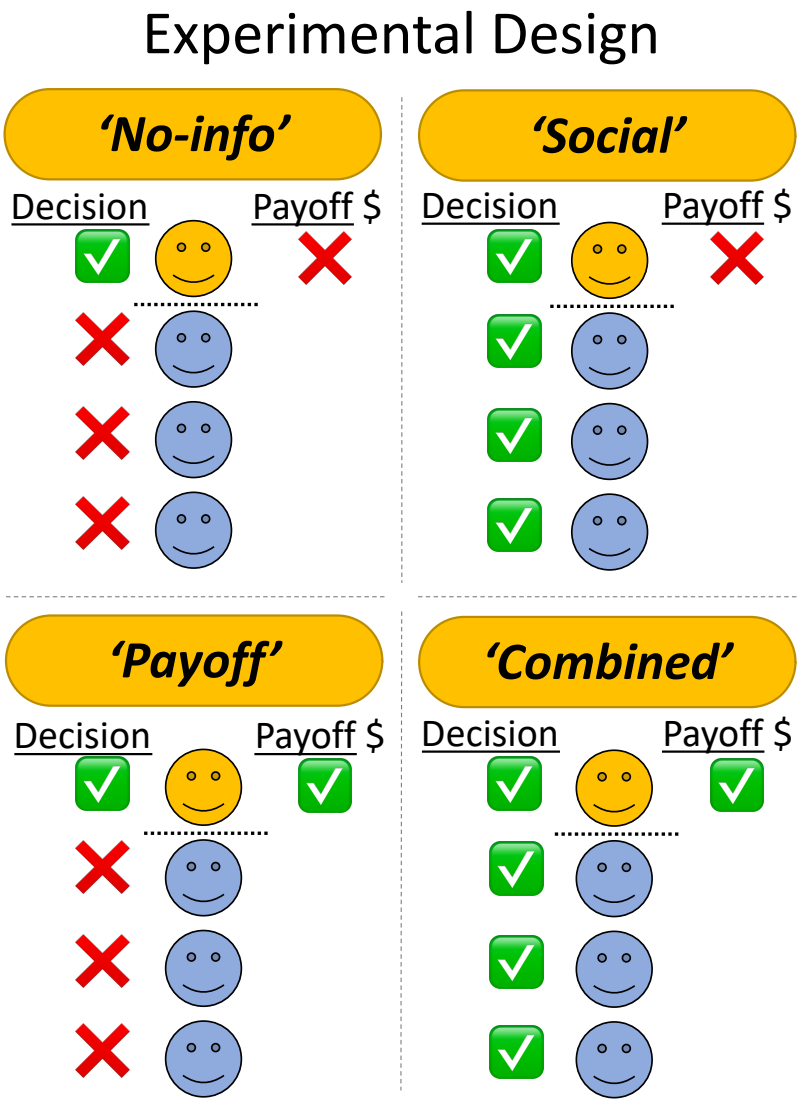


Figure 1. Experimental design. *We varied how much information participants saw after each round of the game. We decoupled two typical forms of information, either about the focal individual's own earnings (yellow/top individual, payoff information), or the contribution decisions of the focal player's groupmates (blue individuals, social information). This made four general treatments: one treatment with neither form of information ('No-info', N = 21 groups); one with just payoff information ('Payoff', N = 40 groups); one with just social information ('Social', N = 53 groups; 33 groups received only limited information, the group average); and one with both forms of information ('Combined', N = 40 groups; 20 with only limited social information, the group average). All individuals within the same group received the same treatment. We never showed the earnings (payoff) of groupmates.*

METHODS

Data collection

We conducted two studies, a year apart, identical in design except that degree of social information showed either all individual contributions (full social information), or just the group average contribution (limited social information). In both cases adding payoff information is technically redundant because both forms can be used to calculate personal payoffs, but only the limited social information can be calculated from personal payoffs. Study 1, with detailed social information used 280 participants (70 groups of four) in the Autumn semester of 2018, from 21st to 30th November, and Study 2, with just the group average, used 336 participants (84 groups of four) in the Autumn semester of 2019, from 7th to 11th October. All data were anonymous and collected electronically. Our experimental, data and analysis files are available online:
[https://osf.io/t4smj/?view_only=a7e682a10a7d45dda6d463d6cd65c402]

Participants and Location

In total we ran 40 sessions with 616 participants. Each session involved 12-20 participants (3-5 groups). All sessions were conducted at the Faculty of Business and Economics (HEC), University of Lausanne (UNIL), Switzerland, in the HEC-LABEX facility. LABEX forbids deceiving participants and requires that all experimental designs obtain prior ethical approval from the LABEX ethics committee. We recruited our participants from the HEC-LABEX participant pool using ORSEE and excluded all participants from previous experiments by the same authors (Greiner, 2015). Participants were mostly students enrolled at either UNIL or the Swiss Federal Polytechnic School (EPFL), and they can be from diverse ethnic, socio-economic, and cultural backgrounds (these variables were not recorded). According to the questionnaires we had a near equal gender ratio (306 Female, 304 Male, 2 Other, and 4 declined to answer) and most of our participants were under 26 years of age (277 aged under 20, 309 aged 20-25, 25 aged 26-30, 2 aged 30-35, 2 x Over 35, and 1 declined to answer, responses to the age question were categorical to increase anonymity). In study 1 we excluded HEC students. In study 2 we allowed first year students from HEC.

Financial incentives

We paid all participants a show up fee of 10 CHF. For each of the 11 incentivized decisions, we endowed participants with 20 Monetary Units (MU) and each MU was worth 0.025 CHF, so 20 MU was worth 0.5 CHF.

Experimental procedure

The experiment had several stages (Supplementary Figure 1). We provide a full copy of the English instructions in our Supplementary Methods. For each stage participants had to answer questions regarding their understanding of the instructions ('control questions'). For all questions, participants were allowed unlimited time and two attempts before we showed them the correct answers. The initial instructions often took a very long time (more than 20 minutes) (Bigoni & Dragone, 2012).

Participants were randomly assigned to a partitioned computer terminal. We announced that communication was forbidden, and that the HEC-LABEX forbids deception. Participants then progressed, at their own pace, through on-screen instructions and control questions based on hypothetical scenarios detailing the general public-goods game decision and payoffs. We used publicly available instructions, control questions, and structural parameters and translated them into French (U. Fischbacher & Gachter, 2010). The specific parameters were group sizes of four participants, an endowment of 20 Monetary Units, and a Multiplier/Efficiency Factor of 1.6, which meant all contributions were multiplied by 1.6 before being shared out equally, to give a marginal per capita return (MPCR) of 0.4. As this is less than 1, the rational selfish income maximizing decision in a single round is to contribute 0 MU. At this stage the instructions have only described the general decision task and the payoffs involved. The instructions so far have not mentioned game length, type of groupmates or information feedback.

We then used an asocial control treatment to test if individuals maximized their income when there were no social concerns (pre-game 'Income Maximization Test'). We informed participants that they would face the same decision but in 'a special case' with computerized groupmates programmed to play randomly. We assured them that this decision would not be seen by other participants, would have no future consequences, and no financial consequences for other participants but did have financial consequences for

them. We made them answer five true or false statements about the situation they faced and provided them with the correct answers afterwards (Supplementary Methods). After their decision, we did not provide them with any immediate feedback about their earnings or the computer decisions, to prevent learning about the game before the information treatments. This income maximization test allowed us to estimate how much individuals still contribute even when they can have no rational concern for fairness or helping others.

Next, participants learned that they would now face the same decision but with real people, as outlined in the original instructions. We informed them that they had been randomly grouped with three other participants for nine rounds of decision making. Although repeated games can favor some strategic cooperation, backwards induction would mean that the selfish income maximizing decision in a finite game of known length among rational individuals is still to contribute 0 MU (Ambrus & Pathak, 2011; M. N. Burton-Chellew, C. El Mouden, & S A. West, 2017a; Dijkstra & van Assen, 2017; Kreps, Milgrom, Roberts, & Wilson, 1982; Krockow, Colman, & Pulford, 2016). We then explained to each group, according to its randomly assigned treatment, what type of information they would all be shown after each round. We assured them that all groupmembers would receive the same type of information. Again, we also made them answer five true or false statements about the situation they faced.

Specifically, we told participants, in the:

No-info treatment, "*You and everyone else in the group will not receive any information after each round. No participants will be able to know your investments at any time. Your earnings will not be shown to you each round, but you will receive the money at the end of the experiment.*".

Social treatment, "*The information that each person will receive will only be the decision of each group member / average decision of the group* [study 1 / study 2]. *Your earnings will not be shown to you each round, but you will receive the money at the end of the experiment*."

Payoff treatment, "*The information that each person will receive will only be their own earnings in each round. No participants will be able to know your investments at any time.*"

Combined treatment, "*The information that each person will receive will only be their own earnings, and the decision of each group member / and the average decision of the group* [study 1/ study 2]."

After the repeated public goods game with information treatments the participants were then told that they would face the same decision again but again in the special case with computers for just one round (post-game Income Maximization Test). We again assured them that this decision would have no future consequences but did have financial consequences for them and no other participants. The post-game Income Maximization Test allowed us to measure how the different information treatments affected learning about how to maximize income.

At the end of the experiment, we asked participants two unincentivized questions regarding their understanding of the game's payoffs, and their motivation when grouped with humans. Specifically, to test understanding of the game's payoffs, we asked them, "*In the basic decision situation, played for one round only, if a player wants to maximize his or her earnings, should they decide their contribution depending on what the other people in their group contribute?*". Participants had a choice of four answers: the correct answer, "*No*"; or two incorrect answers; "*Yes*", "*Sometimes*", or they could respond "*Do not know*" (Burton-Chellew, et al., 2016). Asking participants about a strictly one-shot game allowed us to more cleanly measure their understanding of the game's payoffs and the social dilemma involved than if we had asked them about a repeated game. If a participant does not know the correct answer to this question, then they do not understand the game's payoffs.

To measure motivations we asked, "*Which of these descriptions best describes your motivation during the rounds with humans?*" and offered participants four possible responses, that corresponded to selfishness ("*Making myself as much money as possible*."); or competitiveness ("*Making myself more money than other people*."); or a desire for fairness ("*Avoiding unequal outcomes so that I make neither more, nor less, than other people*."); or a desire to help the group ("*Making the group as much money as possible even if it meant making myself less money.*"). For both questions we reversed the on-screen order of responses for half of the participants (2 within each group).

Statistical analyses

We analyzed the data using R-Studio version 1.3.1093 (Team, 2020). All statistical tests were two-tailed. Significance values in linear mixed models were estimated by the lmerTest package, which uses the Satterthwaite approximation for estimating degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2017). Data figures were made with the ggplot2 package (Wickham, 2009).

RESULTS

Payoff information led to a greater decline than social information. In support of the Confused Learners hypothesis, contributions declined faster when we showed individuals their payoff information instead of social information (Figure 2). Pooling all the data together, we find that the rate of decline was significantly faster among those shown only their payoff compared to those shown just social information (Full model controlling for Study, Payoff versus Social treatment comparisons, lmer: estimated difference = -1.7 percentage points per round, 95% CI = [-0.60, -2.73], $t_{1,154.0}$ = -3.1, P = 0.002, Table 1). Specifically, when individuals saw payoffs, contributions declined by -3.6 percentage points per round, decreasing from 49% to 18% over 9 rounds, whereas among those shown social info, contributions only declined at a rate of -1.9 percentage points per round, from 47% to 29% (Table 1).

Social information had a negligible effect. In contrast, the effect of showing individuals social information had a negligible effect upon the rate of decline. Compared to those individuals who saw no information (No-info treatment), the rate of decline was not significantly faster when individuals saw only information on the decisions of their groupmates (Social treatment) (lmer, Contribution ~ Game Round * Information treatment + Study: estimated difference in decline between No-info and Social treatments = -0.9 percentage points per round, 95% CI = [2.17, -0.45], $t_{1,154.0}$ = -1.3, P = 0.197, Figure 2, Table 1). Repeating the above analyses separately for each study, which had different levels of social information, did not qualitatively change the conclusions (Supplementary Results, Supplementary Figure 2, Supplementary Tables 1 & 2), and there was no overall significant

three-way interaction between study and game round and information treatment (linear mixed model, Study * Game round * Information treatment, $F_{3,154} = 0.7$, $P = 0.575$).

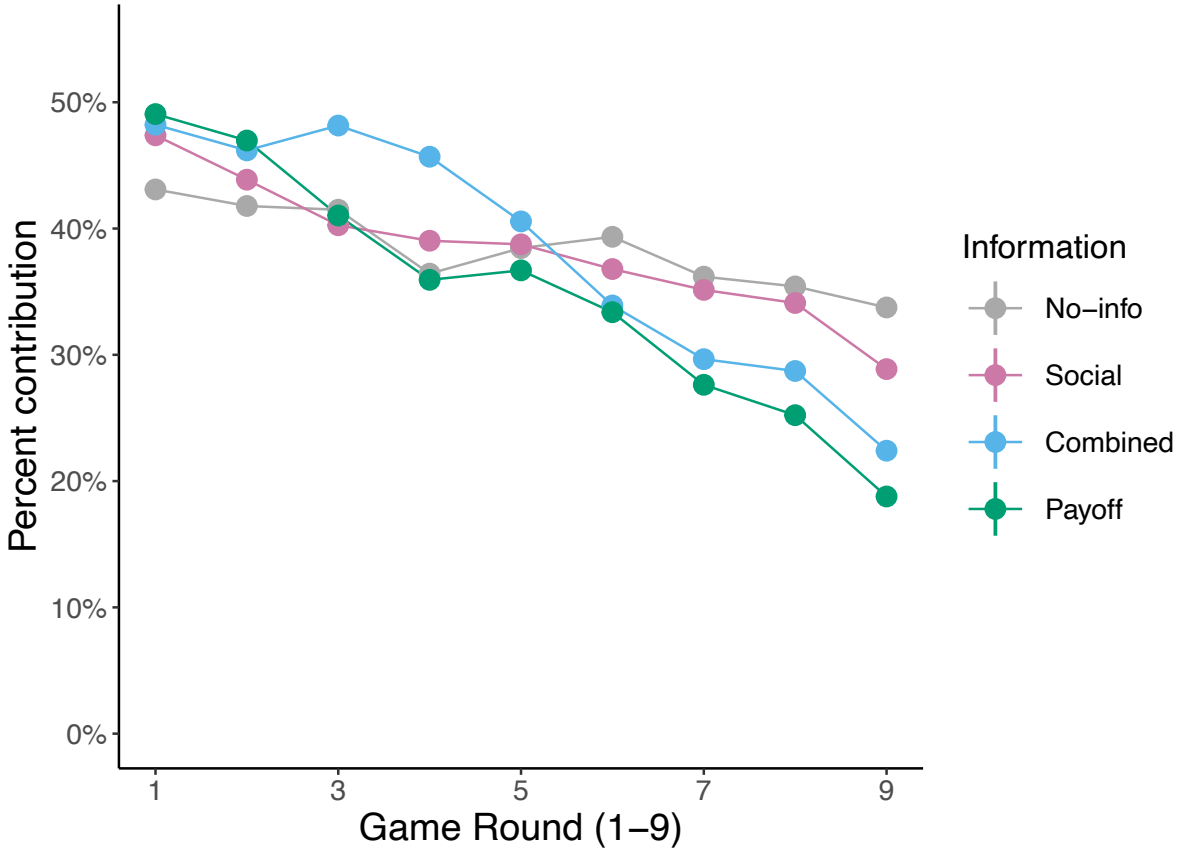


Figure 2. Declining cooperation. *Depending on randomly assigned treatment, groups of individuals were shown either no information (grey, 21 groups); or social information (magenta, 53 groups); or payoff information (green, 40 groups) or both social and payoff information combined (blue, 40 groups). Payoff information led to a greater decline than social information. The data were collected across two studies that varied in the level of social information shown, including either all individual decisions or just the group average (Supplementary Figure 2). This figure is for illustration purposes and does not evaluate random effects.*

Table 1. Declining cooperation. Results from a linear mixed model fit by maximum likelihood on the percent contribution by each group per round depending on information treatment (reference treatment: No-info).

Parameter	Coefficient	Std.Error	t[1]	df	p-value	Sig.
(Intercept) [No-info]	47.73	4.372	10.918	159.2	<0.001	***
Game round [No-info]	-1.07	0.563	-1.9	154.0	0.059	
Year of study [2019]	-9.57	2.803	-3.415	154.0	<0.001	***
Treatment [Social]	4.20	4.874	0.861	154.0	0.391	
Treatment [Payoff]	6.40	5.086	1.259	153.8	0.210	
Treatment [Combined]	8.85	5.086	1.74	153.8	0.084	
Game round*Treatment [Social]	-0.86	0.665	-1.295	154.0	0.197	
Game round*Treatment [Payoff]	-2.53	0.665	-3.635	154.0	<0.001	***
Game round*Treatment [Combined]	-2.34	0.665	-3.363	154.0	<0.001	***

Number of observations = 1,386

Number of independent groups = 154

[1]t-tests use Satterthwaite's method for estimating degrees of freedom.

We confirmed our main result, that the rate of decline was faster when payoff information was available, with two alternative analyses. First, we found, in support of the Confused Learners hypothesis, that the rate of decline was faster among those shown both social and payoff information (Combined information treatment), than among those shown just social information (Social treatment, pooling both studies, the rate of decline = -1.5 percentage points faster per round in the Combined treatment, 95% CI = [-0.41, -2.54], $t_{1,154.0} = -2.7$, P = 0.007). This means that the addition of payoff information to social information led to a faster decline, despite this information being technically redundant if individuals perfectly understand the game. In contrast, the addition of social information to payoff information did not lead to a significantly faster decline in the Combined information treatment (pooling both studies, the rate of decline = 0.2 percentage points slower per round in the Combined treatment than in the Payoff treatment, 95% CI = [1.33, -0.95], $t_{1,154.0} = 0.3$, P = 0.744). In summary, adding payoff information accelerates the decline, but adding social information does not.

Second, we analyzed all the groups separately, and found that groups which saw payoff information were almost twice as likely to develop a negative correlation between their contributions and round of the game (percentage of groups with a significant negative Pearson correlation, Payoff treatment = 70%; Combined treatment = 65%; Social treatment = 36%, No-info treatment = 24%; Number of groups = 40/40/53/21 respectively, Table 2). Among those groups in the Social treatment, 17% even finished at a higher level than their starting level (N = 9/53). In contrast, only 2% of groups in the Payoff treatment finished at a higher level (N = 1/40) (Table 3, Supplementary Figures 3-6 show the time profiles for each group). These comparisons would suggest that the Social treatment, even though it implicitly contains payoff information (personal payoffs can be calculated from the social information), was behaviorally more akin to the No-info treatment than to the two treatments showing explicit payoff information (Payoff and Combined treatments).

Table 2: Pearson correlations. Summary of the correlations between contributions and the nine rounds of the game depending on the information treatment. Pearson correlations calculated for each individual and for each group.

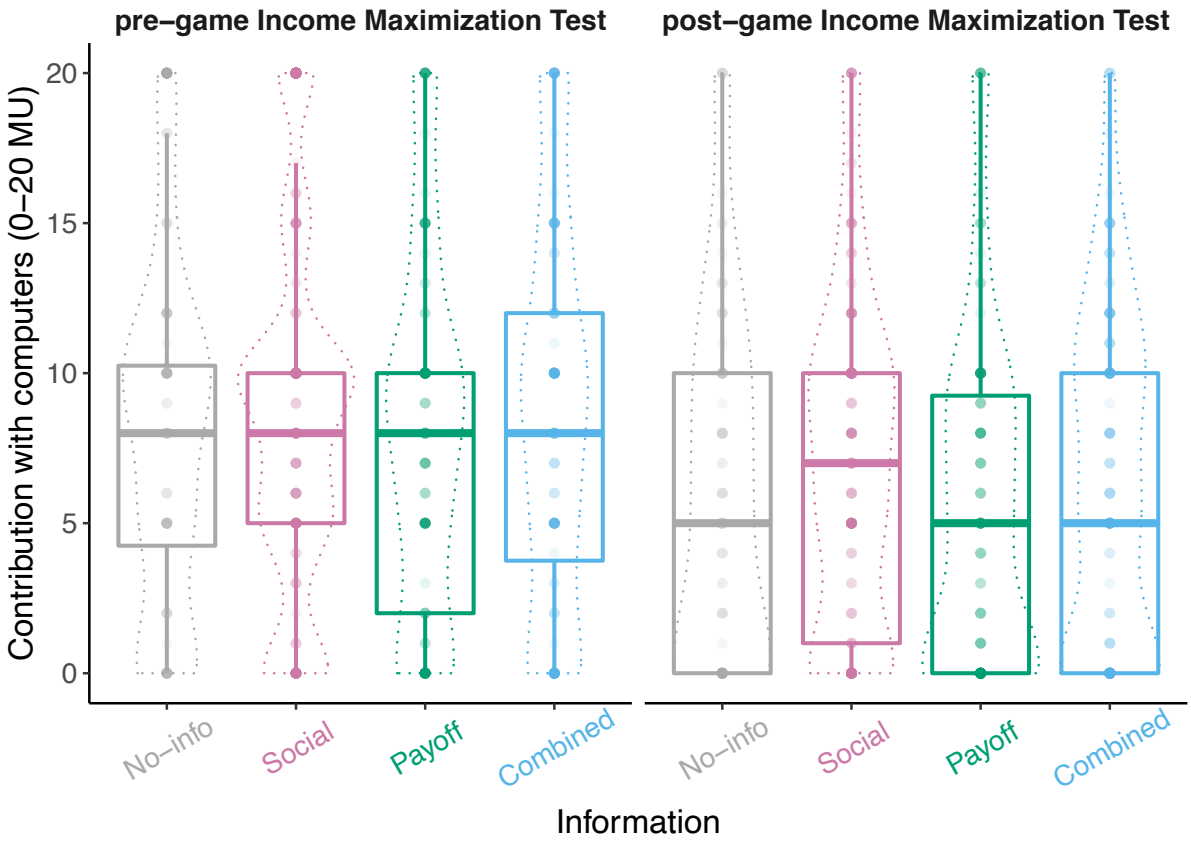
	Individuals					Groups				
Treatment	Mean	Median	Sig. neg.	Sig. pos.	N_i	Mean	Median	Sig. neg.	Sig. pos.	N_g
No-info	-0.12	0.00	10% (8)	1% (1)	84	-0.26	-0.16	24% (5)	0% (0)	21
Social	-0.22	-0.29	14% (30)	5% (11)	212	-0.37	-0.55	36% (19)	4% (2)	53
Payoff	-0.40	-0.45	28% (45)	1% (1)	160	-0.70	-0.77	70% (28)	0% (0)	40
Combined	-0.37	-0.44	30% (48)	3% (5)	160	-0.62	-0.77	65% (26)	5% (2)	40

Table 3: Comparing start and finish. Summary of the differences in contributions between the beginning (Round 1) and end (Round 9) of the game by information treatment. Differences calculated for each individual and for each group.

			Individuals					Groups		
Information	Mean (MU)	Median (MU)	N decrease	N increase	N same	N_i	N decrease	N increase	N same	N_g
No-info	-1.9	0	43% (36)	19% (16)	38% (32)	84	76% (16)	24% (5)	0% (0)	21
Social	-3.7	-3	62% (131)	25% (52)	14% (29)	212	83% (44)	17% (9)	0% (0)	53
Payoff	-6.1	-5	68% (109)	11% (17)	21% (34)	160	98% (39)	2% (1)	0% (0)	40
Combined	-5.2	-5	64% (103)	16% (25)	20% (32)	160	90% (10)	10% (4)	0% (0)	40

A range of alternative decision tasks confirmed most participants are confused and self-interested. We also supplemented our above approach with a range of three alternative decision/survey tasks (Methods, Supplementary Figure 1).

Alternative task 1: most participants failed to maximize income in the asocial control. In support of the Confused Learners hypothesis we found (1) that in the pre-game income maximization test (IMT1), which was one round of the same public-goods game but played knowingly with computerized groupmates, most participants behaved irrationally (82% contributed more than 0 MU, N = 504/616; mean average contribution = 8.2/20 MU, median = 8 MU); and (2) that the greatest performance in the post-game test (IMT2) occurred among those who had experienced the treatments containing payoff information (Payoff and Combined treatments versus Social and No-info treatments, generalized linear quasibinomial model controlling for individual's contribution in the pre-game test: $t_{1,612} = -2.6$, P = 0.009, Figure 3, Supplementary Figure 7, Supplementary Table 3).



374 ***Figure 3. Income Maximization Tests.*** *Most individuals failed to maximize their income even*
375 *when grouped with computers. We tested all individuals twice, one before and once after*
376 *the repeated public goods game. The least improvement was among individuals shown only*
377 *social information (pink). Number of individuals: No-info = 84; Social = 212; Payoff =160;*
378 *Combined = 160.*
379
380 **Alternative task 2: most participants appeared confused about payoffs.** We also found
381 evidence of substantial confusion when we asked participants if the income-maximizing
382 strategy depended on the actions of their groupmates or not (Methods) (Burton-Chellew, et
383 al., 2016)? Only 25% of participants answered correctly ('*No*') (N=153/616). In contrast, 72%
384 of participant answered incorrectly (N=444), indicating that they had false beliefs about the
385 game's Payoff, with 53% responding that "*Yes*", an income maximizing individual should
386 decide their contribution depending on what others contribute (N=325) and another 19%
387 responding "*Sometimes*" (N=119). The remaining participants, 3%, indicated that they did
388 not know ("*Do not know*", N = 19) (Table 2). Incorrect individuals were significantly more
389 likely to have cooperated with computers (numbers that contributed more than 0 MU
390 towards computers: in IMT1, N = 399 of 463 non-correct responses versus 105 of 153
391 correct responses; Fisher's Exact Test: P < 0.001; in IMT2, N = 358/463 versus 75/153;
392 Fisher's Exact Test: P < 0.001).
393

Table 4: Confusion about Payoffs. Summary of participant responses to: "*In the basic decision situation, played for one round only, if a player wants to maximize his or her earnings, should they decide their contribution depending on what the other people in their group contribute?*"^

Information	"Yes"	"No"*	"Sometimes"	"Do not know"	N
No-info	54% (45)	25% (21)	20% (17)	1% (1)	84
Social	50% (105)	26% (55)	21% (45)	3% (7)	212
Payoff	55% (88)	26% (42)	15% (24)	4% (6)	160
Combined	54% (87)	22% (35)	21% (33)	3% (5)	160
Overall	53% (325)	25% (153)	19% (119)	3% (19)	616

^We reversed the order of responses for half of the participants (2 within each group).

* This is the correct response.

394

Alternative task 3: most participants reported selfish motives. Finally, because one important aim of laboratory experiments is to gain knowledge about the specific motives behind altruistic acts (Fehr & Fischbacher, 2003), we asked participants to identify with one of four possible motivations (Methods, Table 3). The results contradicted the idea that most people are fair-minded conditional cooperators. Instead, most participants, 54%, responded that they were best described by the selfish motivation (N = 305/568 surveyed). A further 6% responded that they had a competitive motivation (N = 35), meaning that 60% of participants freely admitted to a socially undesirable response (N = 340). In contrast, only 28% responded that they were best described by the fair or inequity-averse motivation (N = 160) and only 12% that they were motivated by maximizing the group's success even at personal cost (pro-group motivation) (N=68) (Table 3). These reported motivations appeared relatively reliable because they correlated with relative levels of contributions in the game (mean contribution for each motivation was: Competitive = 5.7/20 MU, 95% CI = [4.2, 7.2]; Selfish = 6.5/20 MU, 95% CI = [6.0, 7.0]; Inequity-averse = 8.1/20 MU, 95% CI = [7.4, 8.8]; Pro-group = 11.3/20 MU, 95% CI = [10.2, 12.4]; Linear model: Mean individual contribution over 9 game rounds ~ Motivation, $F_{3,564} = 24.4$, $P < 0.001$).

Table 5: Selfish motivations. Summary of participant responses to: Which of these descriptions best describes your motivation during the rounds with humans?^

Information	Selfish [1]	Inequity-averse [2]	Pro-group [3]	Competitive [4]	N*
No-info	57% (41)	24% (17)	11% (8)	8% (6)	72
Social	51.5% (103)	29% (58)	14.5% (29)	5% (10)	200
Payoff	59% (87)	27% (40)	9% (13)	5% (8)	148
Combined	50% (74)	30% (45)	12% (18)	7% (11)	148
Overall	54% (305)	28% (160)	12% (68)	6% (35)	568

^ We reversed the order of responses for half of the participants (2 within each group).

[1] "Making myself as much money as possible."

[2] "Avoiding unequal outcomes so that I make neither more, nor less, than other people."

[3] "Making the group as much money as possible even if it meant making myself less money."

[4] "Making myself more money than other people."

* 48 participants, 12 in each treatment, were not presented with this question.

DISCUSSION

We tested competing explanations for behavior in public goods games by controlling whether individuals could learn from their own payoffs or not (Confused Learner's hypothesis), and whether they could respond conditionally to their groupmates' contributions or not (Conditional Cooperation hypothesis) (Figure 1). We found substantial support for the Confused Learner's hypothesis across multiple results. Specifically; we found that (1) payoff information generated the greatest decline in contributions, regardless of whether social information was present or not (Figure 2, Tables 1-3) (2) most individuals demonstrated confusion about the game's payoffs when playing with computerized groupmates (Figure 3) or when directly asked (Table 4); and (3) that most individuals admitted to being selfishly motivated to make themselves as much money as possible (Table 5).

Overall, our results suggest that apparent altruism in public goods games mostly arises from confused but self-interested individuals trying to learn how to play the game, and not from fair-minded altruistic individuals trying to help their group or equalize payoffs. If individuals are confused, we cannot be sure what 'game' they think they are playing (Chou, McConnell, Nagel, & Plott, 2009; Columbus, Munich, & Gerpott, 2020; Ferraro & Vossler, 2010; Plott & Zeiler, 2005). For example, if individuals think the best action depends on what others do, like a 'stag-hunt' game or 'threshold' public-goods game (Croson & Marks, 2000; Rondeau, Poe, & Schulze, 2005), then it makes sense for them to act conditionally on social information. This could explain why the frequencies of different responses to our question on how to maximize incomes closely resembled the frequencies of different 'social types' ('Yes' = 'Conditional Cooperator', 'No' = 'Free Rider', etc.) typically reported in other experiments (Burton-Chellew, et al., 2016; U. Fischbacher & Gachter, 2010; U. Fischbacher, et al., 2001; Thoni & Volk, 2018). In this case social types would really be artefacts of variation in levels of understanding (Burton-Chellew, et al., 2016) and repeated public-goods games with payoff information would be measuring rates of learning rather than social preferences (Burton-Chellew & West, 2021).

Our study shows the value of using a range of treatments and methods. For example, our No-info treatment provided an informative baseline treatment when comparing rates of decline (Supplementary Tables 1-2). Our simple, direct, surveys of understanding and motivation complimented our use of the more traditional methods of incentivized decision making ('behavioral economics'). The latter is less susceptible to socially desirable responses, but the former is simpler to understand and the results easier to interpret. Our income maximization tests with computerized groupmates provided quick and simple tests of 'rational' behavior, a commonly assumed hypothesis in economic experiments aiming to measure social behaviors.

Future directions. It could be useful to expand our experiment. One could test which information is of more interest to participants by allowing them to choose (Burton-Chellew & D'Amico, 2021), or to use eye or mouse tracking software (Geran & Weixing, 2020; Jiang, Potters, & Funaki, 2016; Lahey & Oxley, 2016). One could vary how groups are formed, group size, or the length of the game, to investigate if the salience of social information depends on the probability of future social interactions (Burton-Chellew, et al., 2017a; Fiala & Suetens, 2017; Reuben & Suetens, 2012; Trivers, 2006). One could ask participants about their expectations about their groupmates' contributions to gain insights into their thought processes (Chaudhuri, Paichayontvijit, & Smith, 2017; U. Fischbacher & Gachter, 2010). We did not do so because asking individuals to guess the contributions of their groupmates may distort behavior by stimulating conditional cooperation in individuals that would otherwise not have thought about their groupmates. One could use a range of different instructions to test if our results generalize to different set-ups or different cultures (Li, 2017). Here we simply replicated 'standard' instructions and used a common participant pool, to enable comparisons with many prior key studies (Fehr & Gachter, 2002; U. Fischbacher & Gachter, 2010; U. Fischbacher, et al., 2001). We replicated the standard result, but our expanded design and varied approach allowed us to identify that most participants were in fact selfish, not altruistic, but needed to learn about the game's payoffs.

ACKNOWLEDGEMENTS

We are grateful to Tiago Fernandes for assistance with translations and conducting the first study; Luis Santos-Pinto and Laurent Lehmann for helpful discussion, and the audience of

the Behavioral Economics and Experimental Research group at UNIL for feedback. Funding provided by Professor Laurent Lehmann's independent research budget.

References

Ambrus, A., & Pathak, P. A. (2011). Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics, 95*, 500-512.

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

Andreozzi, L., Ploner, M., & Saral, A. S. (2020). The stability of conditional cooperation: beliefs alone cannot explain the decline of cooperation in social dilemmas. *Scientific Reports, 10*, 13610.

Angelovski, A., Di Cagno, D., Güth, W., Marazzi, F., & Panaccione, L. (2018). Behavioral spillovers in local public good provision: An experimental study. *Journal of Economic Psychology, 67*, 116-134.

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

Bigoni, M., & Dragone, D. (2012). Effective and efficient experimental instructions. *Economics Letters, 117*, 460-463.

Burton-Chellew, M. N., & D'Amico, V. (2021). Humans prefer to copy success rather than the social norm in a cooperative game. *SocArXiv*.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 1291-1296.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2017a). Evidence for strategic cooperation in humans. *Proceedings of the Royal Society B: Biological Sciences, 284*.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2017b). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society B-Biological Sciences, 284*.

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff based learning explains the decline in cooperation in human public goods games. *Proceedings of the Royal Society B-Biological Sciences, 282*, 20142678.

Burton-Chellew, M. N., & West, S. A. (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 216-221.

Burton-Chellew, M. N., & West, S. A. (2021). Payoff-based learning best explains the rate of decline in cooperation across 237 public-goods games. *Nature Human Behaviour*.

Camerer, C. F. (2013). Experimental, cultural, and neural evidence of deliberate prosociality. *Trends in Cognitive Sciences, 17*, 106-108.

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

Chaudhuri, A., Paichayontvijit, T., & Smith, A. (2017). Belief heterogeneity and contributions decay among conditional cooperators in public goods games. *Journal of Economic Psychology, 58*, 15-30.

Chou, E., McConnell, M., Nagel, R., & Plott, C. R. (2009). The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person "guessing" game. *Experimental Economics, 12*, 159-179.

Columbus, S., Munich, J., & Gerpott, F. H. (2020). Playing a different game: Situation perception mediates framing effects on cooperative behaviour. *Journal of Experimental Social Psychology, 90*.

Cooper, D. J., & Stockman, C. K. (2002). Fairness and learning: an experimental examination. *Games and Economic Behavior, 41*, 26-45.

Croson, R., Fatas, E., & Neugebauer, T. (2005). Reciprocity, matching and conditional cooperation in two public goods games. *Economics Letters, 87*, 95-101.

Croson, R., & Marks, M. (2000). Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. *Experimental Economics, 2*, 239-259.

Di Cagno, D., Galliera, A., Güth, W., & Panaccione, L. (2016). A hybrid public good experiment eliciting multi-dimensional choice data. *Journal of Economic Psychology, 56*, 20-38.

Dijkstra, J., & van Assen, M. A. L. M. (2017). Explaining cooperation in the finitely repeated simultaneous and sequential prisoner's dilemma game under incomplete and complete information. *Journal of Mathematical Sociology, 41*, 1-25.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*, 785-791.

Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature, 415*, 137-140.

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour, 2*, 458-468.

Ferraro, P. J., & Vossler, C. A. (2010). The Source and Significance of Confusion in Public Goods Experiments. *B E Journal of Economic Analysis & Policy, 10*.

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

Fischbacher, U., & Gachter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review, 100*, 541-556.

Fischbacher, U., Gachter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*, 397-404.

Fischbacher, U., Gächter, S., & Quercia, S. (2012). The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology, 33*, 897-913.

Geran, T., & Weixing, W. (2020). A Review of Mouse-Tracking Applications in Economic Studies. *Journal of Economics and Behavioral Studies, 11*.

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

Gunnthorsdottir, A., Houser, D., & McCabe, K. (2007). Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior & Organization, 62*, 304-315.

Henrich, J., & Muthukrishna, M. (2021). The Origins and Psychology of Human Cooperation. *Annual Review of Psychology, Vol 72, 72*, 207-240.

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

Jiang, T., Potters, J., & Funaki, Y. (2016). Eye-tracking Social Preferences. *Journal of Behavioral Decision Making, 29*, 157-168.

Jouxtel, J. (2019). Voluntary contributions of time: Time-based incentives in a linear public goods game. *Journal of Economic Psychology, 75*, 102139.

Kocher, M. G., Martinsson, P., Persson, E., & Wang, X. (2016). Is there a hidden cost of imposing a minimum contribution level for public good contributions? *Journal of Economic Psychology, 56*, 74-84.

Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational Cooperation in the Finitely Repeated Prisoners-Dilemma. *Journal of Economic Theory, 27*, 245-252.

Krockow, E. M., Colman, A. M., & Pulford, B. D. (2016). Cooperation in repeated interactions: A systematic review of Centipede game experiments, 1992-2016. *European Review of Social Psychology, 27*, 231-282.

Kuemmerli, R., Burton-Chellew, M. N., Ross-Gillespie, A., & West, S. A. (2010). Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 10125-10130.

Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 1803-1807.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*, 1-26.

Lahey, J. N., & Oxley, D. (2016). The Power of Eye Tracking in Economics Experiments. *American Economic Review, 106*, 309-313.

Ledyard, J. (1995). Public goods: a survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of Experimental Economics* (pp. 253-279). Princeton: Princeton University Press.

Li, K. K. (2017). How does language affect decision-making in social interactions and decision biases? *Journal of Economic Psychology, 61*, 15-28.

Nax, H. H., Burton-Chellew, M. N., West, S. A., & Young, H. P. (2016). Learning in a black box. *Journal of Economic Behavior & Organization, 127*, 1-15.

Neugebauer, T., Perote, J., Schmidt, U., & Loos, M. (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology, 30*, 52-60.

Plott, C. R., & Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review, 95*, 530-545.

Reuben, E., & Suetens, S. (2012). Revisiting strategic versus non-strategic cooperation. *Experimental Economics, 15*, 24-43.

Rondeau, D., Poe, G. L., & Schulze, W. D. (2005). VCM or PPM? A comparison of the performance of two voluntary public goods mechanisms. *Journal of Public Economics, 89*, 1581-1592.

Shapiro, D. A. (2009). The role of utility interdependence in public good experiments. *International Journal of Game Theory, 38*, 81-106.

Team, R. (2020). *RStudio: Integrated Development Environment for R*. Boston, MA: http://www.rstudio.com.

Thielmann, I., Böhm, R., Ott, M., & Hilbig, B. E. (2021). Economic Games: An Introduction and Guide for Research. *Collabra: Psychology, 7*.

Thoni, C., & Volk, S. (2018). Conditional cooperation: Review and refinement. *Economic Letters, 171*, 37-40.

Trivers, R. L. (2006). Reciprocal altruism: 30 years later. In P. M. Kappeler & C. P. van Schaik (Eds.), *Cooperation in Primates and Humans: Mechanisms and Evolution*. Berlin: Springer-Verlag.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

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

SUPPLEMENTARY INFORMATION

Contents

SUPPLEMENTARY METHODS

A full copy of the instructions translated into English are shown at the end of these Supplementary Methods. All our experimental files, data and analysis files are available online: Our experimental, data and analysis files are available online: [https://osf.io/t4smj/?view_only=a7e682a10a7d45dda6d463d6cd65c402]

Copy of instructions

Here is a copy of the instructions, which were translated into French for our participants. The instructions for the general game were copied as much as possible from reference (U. Fischbacher & Gachter, 2010)

Stage 1 screen 1 – general instructions

You are now taking part in an economic experiment, which has been financed purely for academic research purposes.
If you read the following instructions carefully, you can, depending on your decisions, earn a reasonable amount of money.
It is therefore very important that you read these instructions with care.

The instructions, which we have distributed to you, are solely for your private information.
You are not allowed to communicate during the experiment.
If you have any questions, please ask us. Violation of this rule will lead to your exclusion from the experiment and all payments.

If you have questions, please raise your hand. A member of the experimenter team will come to you and answer them in private.

During the experiment we shall not speak of CHF but rather of MU (Monetary Units).
During the experiment your entire earnings will be calculated in MU.
At the end of the experiment the total amount of MU you have earned will be converted to CHF at the following rate:
1 MU = 0.05 CHF, so 20 MU = 1 CHF and 100 MU = 5 CHF.
At the end of the experiment your entire earnings from the experiment will be immediately paid to you in cash.

We describe the experimental process below.

Stage 1 screen 2

The decision situation

You will learn how the experiment will be conducted later. We first introduce you to the basic decision situation. You will find control questions at the end of the instructions that help you to understand the decision situation.

You will be in a group consisting of **4 people**. Each group member has to decide on the allocation of 20 MU. You can put these 20 MU into your **private account** or you can invest them **fully or partially** into a project. Each point you do not invest into the project, will automatically remain in your private account.

Your income from the private account:

You will earn one MU for each MU that you put into your private account.
For example, if you put 20 MU into your private account (and therefore do not invest into the project) your income will amount to exactly 20 MU out of your private account.
If you put 6 MU into your private account, your income from this account will be 6 MU.
No one except you earns something from your private account.

Stage 1 screen 3

Your income from the project:

Each member of the group will also benefit from the amount you invest in the project.
On the other hand, you will also gain from the other group members' investments.
The income for each group member will be determined as follows:

Income from the project = *sum of all contributions x 0.4*

If, for example, the sum of all contributions to the project is 60 MU, then you and the other members of your group each earn 60 x 0.4 = 24 MU out of the project.
If four members of the group contribute a total of 10 MU to the project, you and the other members of your group each earn 10 x 0.4 = 4 MU.

Total income:

Your total income is the sum of your income from your private account and that from the project:

Income from your private account (= 20 - contribution to the project)
+ Income from the project (= 0.4 x sum of all contributions to the project)
= Your Total Income

Stage 1 screen 4 – Control Question 1/4, for all questions, participants were allowed unlimited time and two attempts before we showed them the correct answers

Please answer the following four control questions. They will help you to gain an understanding of the calculation of your income, which varies with your decision about how to distribute your 20 MU.

There is a calculator available.

To access the calculator, click on the image in the bottom right of your screen.

Please enter your answers into the spaces provided and press Continue when done.

Question 1: Each group member has an endowment of 20 MU. Nobody (including yourself) contributes any MU to the project.

What is *your* total income (in MU)? [answer box, correct answer is 20]

What is the total income of each of the *other* group members? [answer box, correct answer is 20]

Stage 1 screen 5

Here is the second question.

Question 2: Each group member has an endowment of 20 MU. You invest 20 MU in the project. Each of the other three members of the group also contributes 20 MU to the project.

What is *your* total income (in MU)? [answer box, correct answer is 32]

What is the total income of each of the *other* group members? [answer box, correct answer is 32]

Stage 1 screen 6

Here is the third question.

Question 3: Each group member has an endowment of 20 MU. The other three group members contribute a total of 30 MU to the project.

What is *your* total income (in MU), if you - in addition to the 30 MU - invest 0 MU into the project? [answer box, correct answer is 32]

What is *your* total income (in MU), if you - in addition to the 30 MU - invest 10 MU into the project? [answer box, correct answer is 26]

What is *your* total income (in MU), if you - in addition to the 30 MU - invest 15 MU into the project? [answer box, correct answer is 23]

Stage 1 screen 7

Here is the final question.

Question 4: Each group member has an endowment of 20 MU. Assume that you invest 8 MU to the project.

What is *your* total income (in MU) if the other group members together - in addition to your 8 MU - contribute another 7 MU in total to the project? [answer box, correct answer is 18]

What is *your* total income (in MU) if the other group members together - in addition to your 8 MU - contribute another 12 MU in total to the project? [answer box, correct answer is 20]

What is *your* total income (in MU) if the other group members together - in addition to your 8 MU - contribute another 22 MU in total to the project? [answer box, correct answer is 24]

Stage 2 screen 1 – pre-game Income Maximization Test (IMT1)

Before beginning, you are going to play this game in a **special case. You will face the same decision but in a special case.**

In this special case, **your income will be calculated in the same was previously described.**

However, in this special case, you will be in a group of just you and the COMPUTER.

The situation will be the same as described in the instructions,

but instead of 3 other people, it will just be you and **the computer.**

The computer will make the decisions of the other 3 players.
The computer will choose their decisions **randomly and separately** (so each computer player will make its own random decision).

You are the only person in the group, and only you will receive any money.
Nobody but the experimenter will know your decision.

You will now play this special case for only one round.
You will not be told what happened but will receive the money at the end of the experiment. After playing this special case, you will go on to play in groups of 4 people as described in the instructions.

Nobody but the experimenter will know your decision. Your decision will have no consequences for the rest of the experiment.

Stage 2 screen 2– control questions, we allowed two attempts then showed the correct answers (TRUE/FALSE) on the next screen
Please confirm your understanding of the previous instructions. Enter 1 if the statement is True, or 0 if the statement is False

You are in a group with Computers [answer box, correct answer is 1/TRUE]

You are in a group with Humans [answer box, correct answer is 0/FALSE]

People in your group will see your decision [answer box, correct answer is 0/FALSE]

You will see the decisions of people in your group [answer box, correct answer is 0/FALSE]

You will see your payoffs after each round [answer box, correct answer is 0/FALSE]

Stage 2 screen 3 – pre-game IMT contribution screen

Please enter your contribution (between 0 – 20)

Your contribution to the project is [entry box]

Help box

[Contained a full copy of the instructions regarding the decision and payoff mechanism]

Stage 2 screen 4 – pre-game IMT feedback

You received an endowment of 20 MU

You invested in the group project with computers [X] MU

Your final income from this round is being calculated

You will receive the information on your income at the end of the experiment. Please press continue to move onto the next round.

Stage 3 screen 1 – repeated public goods game with information treatments, certain elements of the instructions varied necessarily according to treatment.

You will now play in a group of 4 people as described in the instructions.

You have been put into a randomly formed group of 4 people.

You will play in this new group, with the same 3 other people, for 9 rounds of decision making.

[No-info treatment]

You and everyone else in the group will not receive any information after each round.
No participants will be able to know your investments at any time.
Your earnings will not be shown to you each round, but you will receive the money at the end of the experiment.

[all other information treatments]
You will receive some information after each round.
You and everybody else in your group will receive the same type of information.

[Social treatment, Study 1 with full social information]
The information that each person will receive will only be the decision of each group member.
Your earnings will not be shown to you each round, but you will receive the money at the end of the experiment.

[Social treatment, Study 2 with limited social information]
The information that each person will receive will only be the average decision of the group.
Your earnings will not be shown to you each round, but you will receive the money at the end of the experiment.

[Payoff treatment]
The information that each person will receive will only be their own earnings in each round.
No participants will be able to know your investments at any time.

[Combined treatment, Study 2 with limited social information]
The information that each person will receive will only be their own earnings and the decision of each group member.

[Combined treatment, Study 2 with limited social information]
The information that each person will receive will only be their own earnings and the average decision of the group.

"*I understand I am now playing with real people*" [annotated continue button]

Stage 3 screen 2 – a repeat of the control questions from Stage 2 (IMT1); we allowed two attempts then showed the correct answers (TRUE/FALSE) on the next screen

Please confirm your understanding of the previous instructions. Enter 1 if the statement is True, or 0 if the statement is False

You are in a group with Computers [answer box, correct answer is 0/FALSE]

You are in a group with Humans [answer box, correct answer is 1/TRUE]

People in your group will see your decision [answer box, correct answer depends, shown in Study 1 only]

People in your group will see the average decision [answer box, correct answer depends, shown in Study 2 only]

You will see the decisions of people in your group [answer box, correct answer depends]

You will see your payoffs after each round [answer box, correct answer depends]

Stage 3 screen 3 – repeated public good game contribution screen

Round X of 9

Please enter your contribution (between 0 – 20)

Your contribution to the project is [entry box]

Help box

[Contained a full copy of the instructions regarding the decision and payoff mechanism]

Stage 3 screen 4 – feedback screens, [varied by treatment]

You received an endowment of 20 MU

You invested in the group project [X] MU

Member B invested [X] MU [Social and Combined treatments, study 1]

Member C invested [X] MU [Social and Combined treatments, study 1]

Member D invested [X] MU [Social and Combined treatments, study 1]

The average decision of the group was [X.Y] MU [Social and Combined treatments, study 2, average shown to 1 decimal place]

You received from the group account: [X] MU [Payoff and Combined treatments]

Your final earnings from this round are: [X] MU [Payoff and Combined treatments]

Your gain is: [X-20] MU [Payoff and Combined treatments]

Your earnings for this round are being calculated [No-info and Social treatments]

Stage 4 screen 1 – post-game Income Maximization Test (IMT2)

You will now make decisions again in the **special case from before with the COMPUTER**.

The decision situation will be the same as before,

but now instead of 3 other people, it will just be you and **the computer again**.

As before, the computer will again pick the decisions of the other 3 group members.

The computer will pick their decisions in the same way as before. That is, **randomly and separately** (so each computer 'member' will make its own random decision).

You are the only person in the group, and only you will receive any money.

You will now make this decision in this special case for only one round.

You will not be told what happened but will receive the money at the end of the experiment.

Nobody but the experimenter will know your decision. Your decision will have no consequences for the rest of the experiment.

Stage 4 screens 2-4 repeated the IMT process of control questions, contributing and feedback (Stage 2 screens 2-4)

SUPPLEMENTARY RESULTS

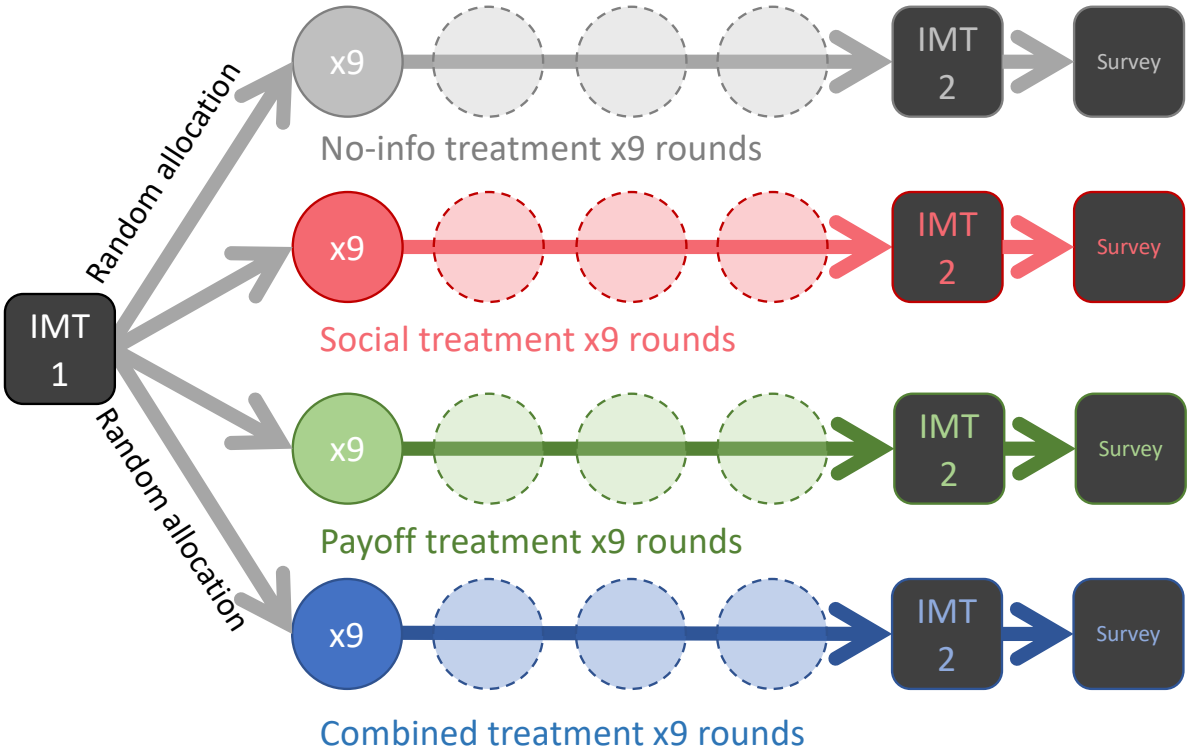
Separate analyses for each study

We found that, in both studies, the rate of decline in cooperation (contributions) significantly depended on the type of information shown (Linear mixed model ran separately for each study, with random intercept and slope for each group, lmer: Group percentage contribution ~ Game round * Information treatment; Study 1, $F_{3,70.0} = 3.1$, P = 0.033; Study 2, $F_{3,84.0} = 5.2$, P = 0.002).

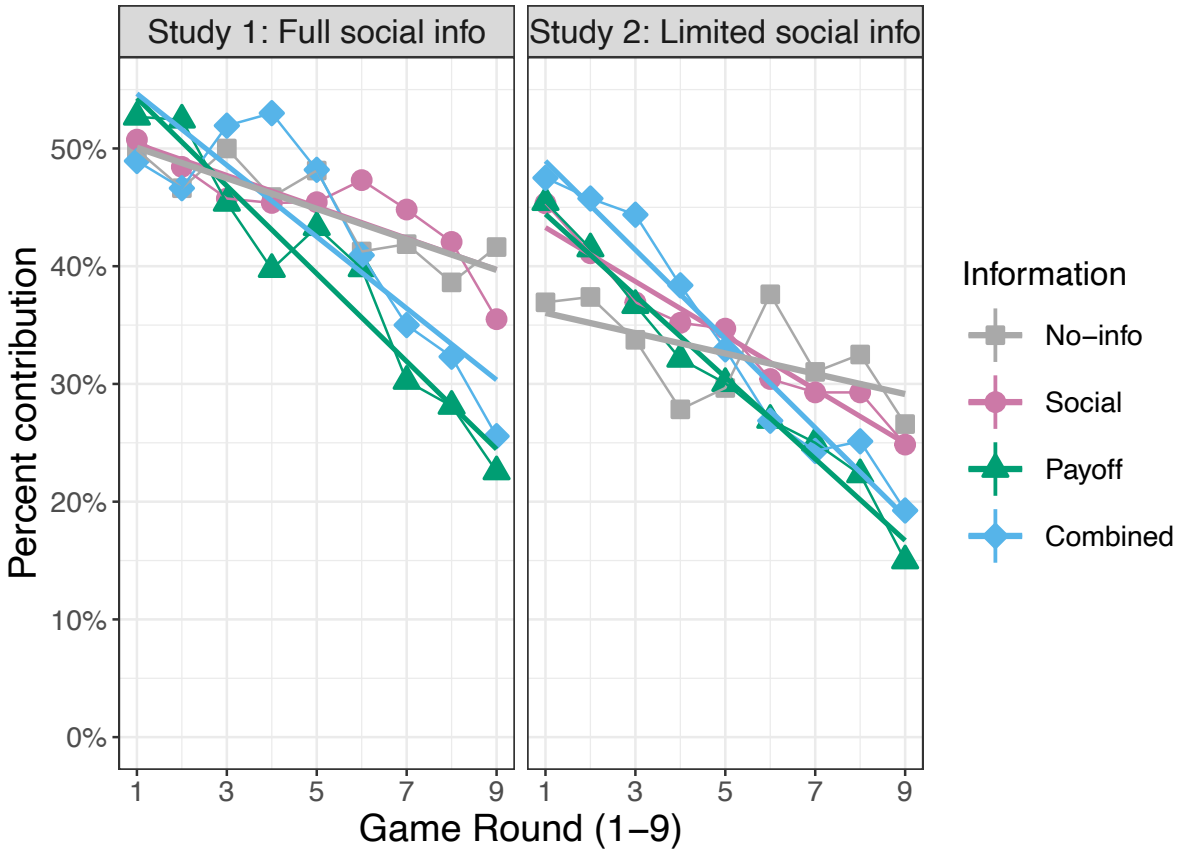
Repeating the analyses from the main Results, but separately for each study, we find that when individuals only saw their payoff and no other information (Payoff treatment), contributions declined at an estimated rate of -3.7 percentage points, 95% CI = [-2.44, -5.02] (Study 1) or -3.5 percentage points, 95% CI = [-2.48, -4.45] (Study 2), per round. This was significantly faster than those shown only detailed social information (Social treatment, Study 1) (lmer: estimated difference between Payoff and Social treatment, Study 1 = -1.7, 95% CI = [-0.57, -4.22], $t_{1,70.0} = 2.6$, P = 0.011, Supplementary Table 1). However the difference was not significant in Study 2 with the limited social information, although the estimate was in the same direction (lmer: estimated difference between Payoff and Social treatment, Study 2 = -1.2, 95% CI = [0.08, -2.42], $t_{1,84.0} = 1.9$, P = 0.067, Supplementary Table 2).

Comparing the effect of adding different levels of social information (individual decisions versus just the group average) to no-information, we found that in both studies, the difference between the Social information treatment and the No-info treatment was non-significant. In study 1, with full social information (shown individual decisions), the estimated difference in the rate of decline was near zero (lmer: estimated difference = -0.03 percentage points per round, 95% CI = [2.20, -2.27], $t_{1,70.0} = -0.03$, P = 0.974, Supplementary Table 1). In Study 2, with limited social information (the group average) there was nearly a

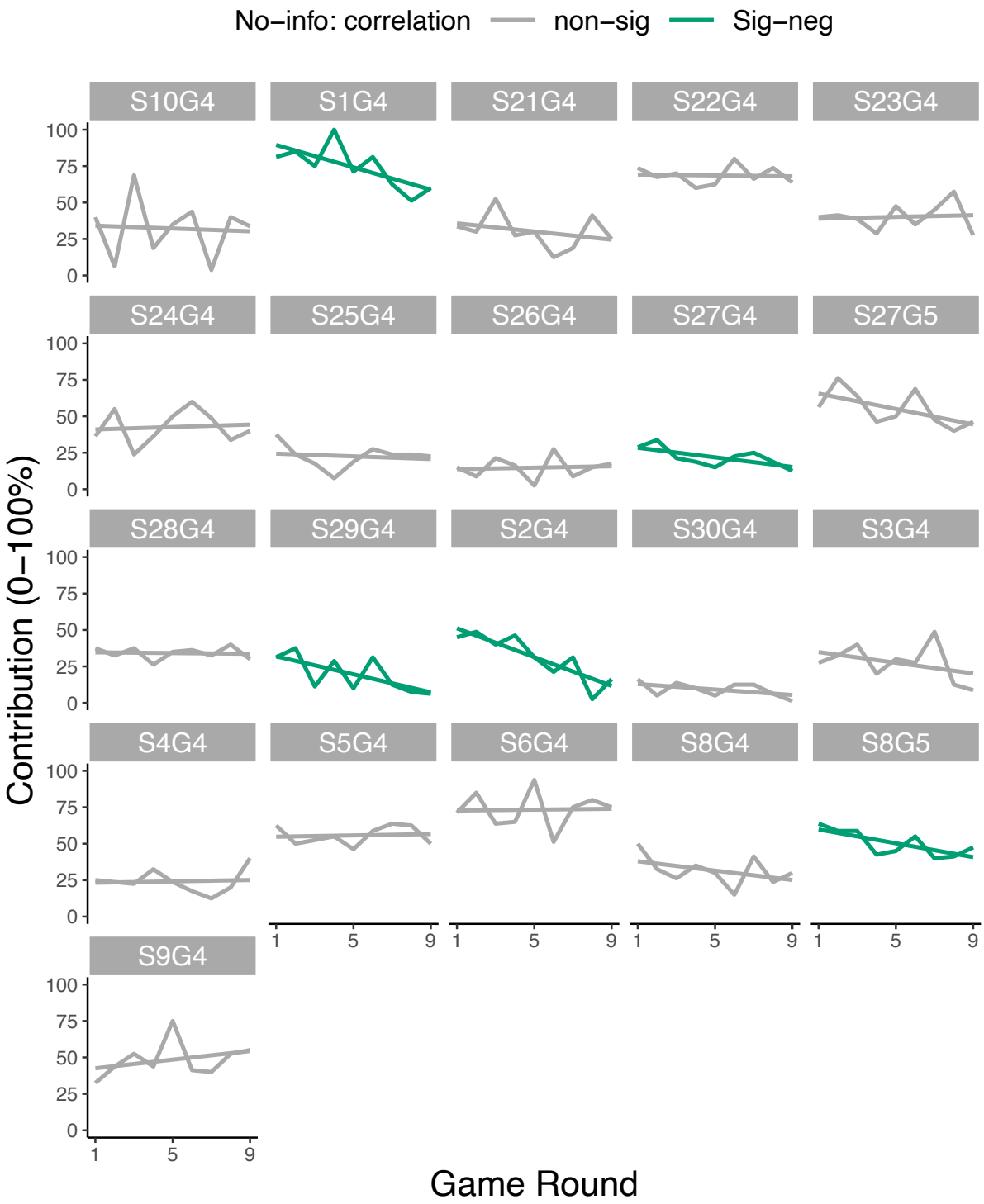
975 significant increase in the decline (lmer: estimated difference = -1.4 percentage points, 95%
976 CI = [0.11, -2.97], $t_{1,84.0}$ = -1.8, P = 0.069, Supplementary Table 2). Therefore it is unclear if
977 the limited social information had an effect (Social treatment, Study 2), because although
978 the slope was almost significantly steeper, the mean level of contribution in the final round
979 was indistinguishable between those that had been shown limited social information, and
980 those that been shown no information at all (Figure 2; linear model on final round group
981 percent contributions: mean contribution in No-info treatment = 26.6%, 95% CI = [16.2,
982 37.0], estimated difference in Social treatment = -1.7% percentage points, 95% CI = [-13.79,
983 10.30], $t_{1,42}$ = -0.3, P = 0.772
984
985 **Comparing rates of decline in the Combined information treatment with the Social and**
986 **Payoff treatments.** The rate of decline in the Combined information treatment which
987 contained both social and payoff information, was significantly faster than in the Social
988 treatment in Study 2 (limited social information) but no in Study 1 (full social information.
989 Specifically, the estimated difference in rate of decline per round between Social and
990 Combined treatments in Study 1 with full social information = -1.7 percentage points, 95% CI
991 = [-3.53, 0.12], $t_{1,70.0}$ = 1.8, P = 0.067; and in Study 2 with limited social information = -1.5
992 percentage points, 95% CI = [-2.73, -0.23], $t_{1,84.0}$ =2.4, P = 0.021, Supplementary Tables 1-2).
993
994 However, the rate of decline in the Combined treatment was not significantly faster than in
995 the Payoff treatment in either study (estimated difference in rate of decline per round
996 between Payoff and Combined treatments in Study 1 with full social information = 0.7
997 percentage points per round, 95% CI = [2.51, -1.14], $t_{1,70.0}$ = 0.7, P = 0.456; in Study 2 with
998 limited social information = -0.3 percentage points per round, 95% CI = [1.09, -1.71], $t_{1,84.0}$ = -
999 0.4, P = 0.661, Supplementary Tables 1-2).
1000
1001



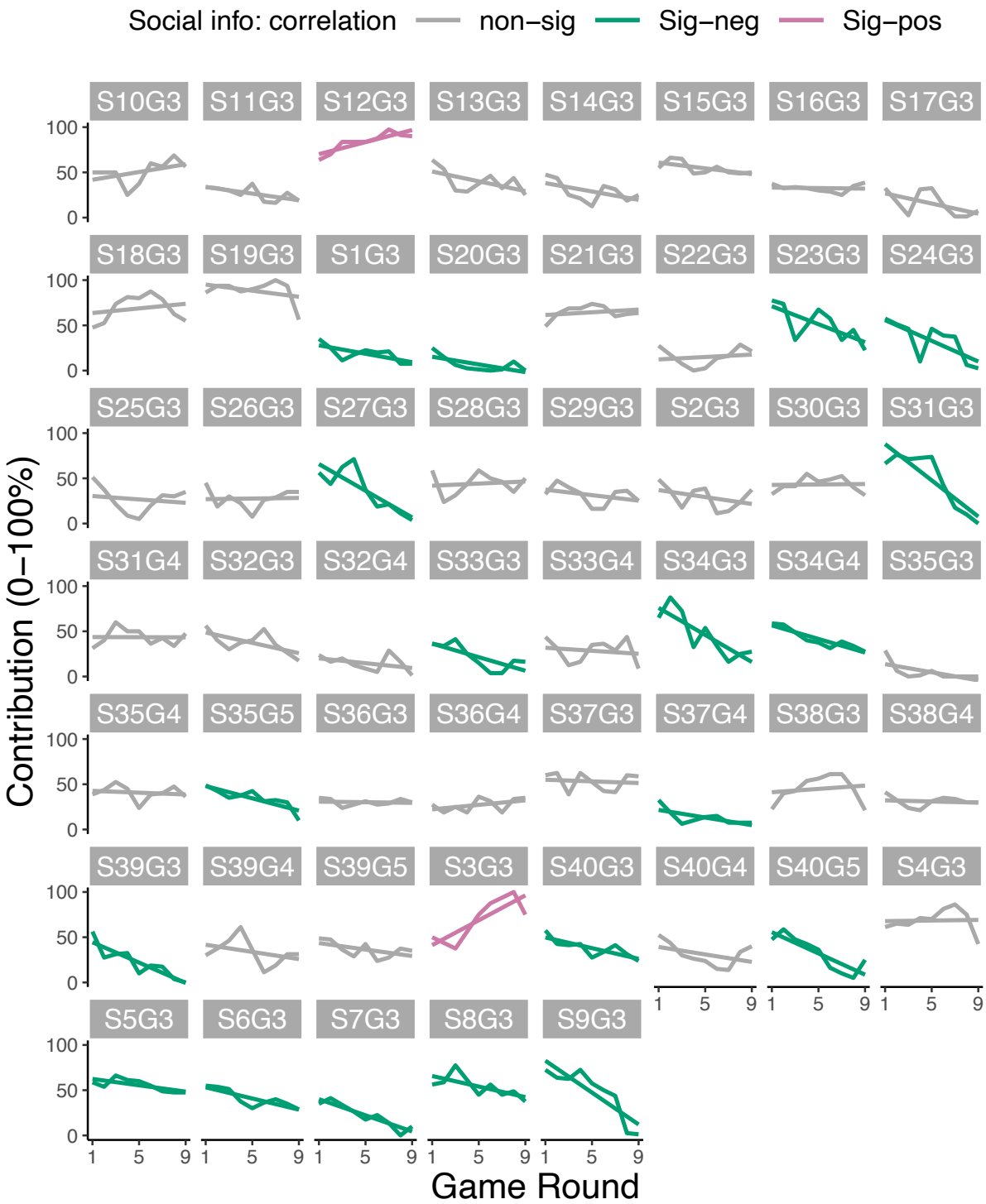
Supplementary Figure 1. Experiment timeline. Participants first had one round of decision making with computerized groupmates in a public-goods game (Income Maximization Test 1). They were then randomly assigned to one of four information treatments and faced nine rounds of the public-goods game in a constant group of four real people. They then repeated the Income Maximization Test (IMT 2) before two survey questions regarding their motivation and their understanding of the public-goods game's payoffs.



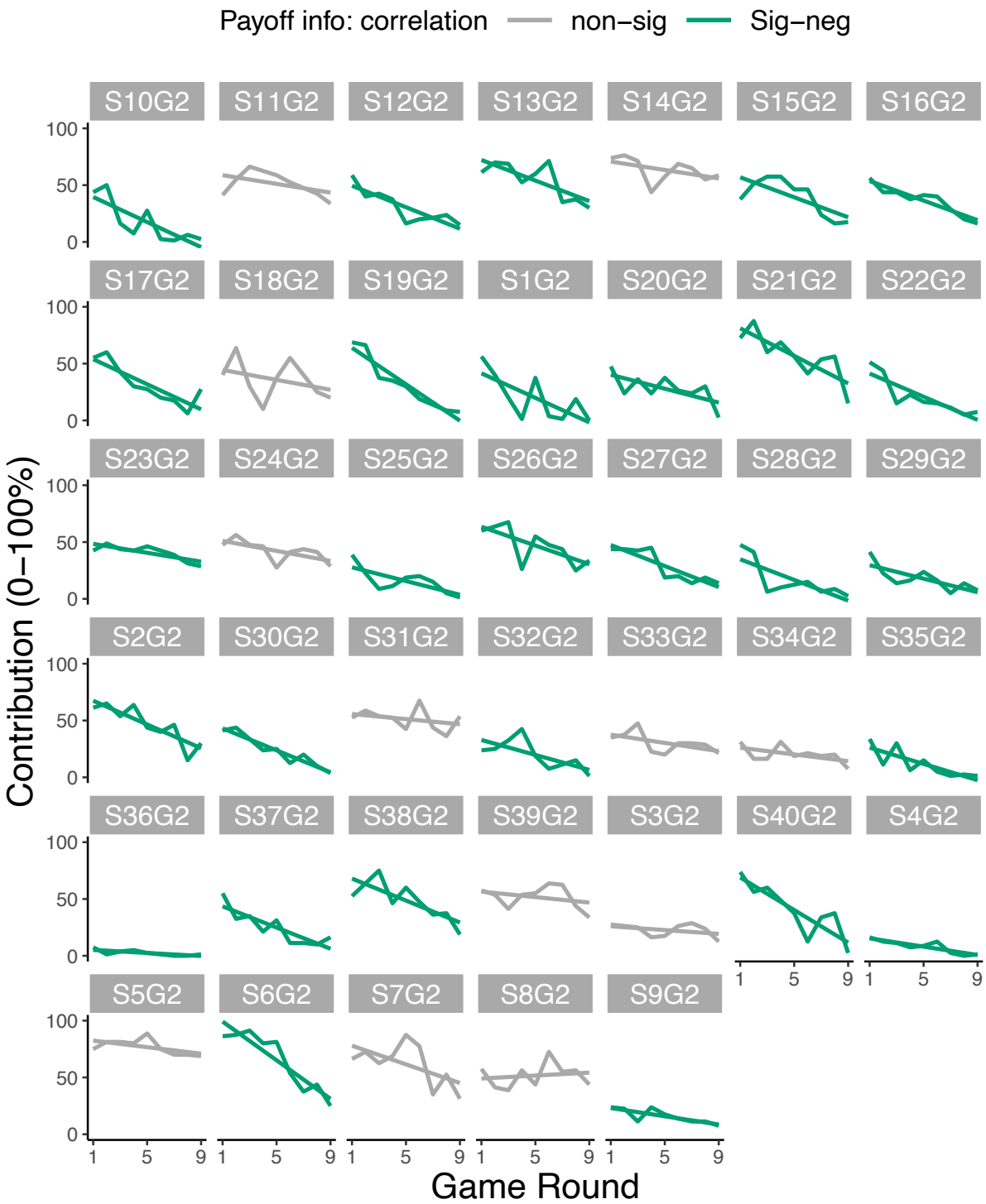
Supplementary Figure 2. Contributions over time for all treatments. This figure is for illustration purposes and does not take into account random effects.



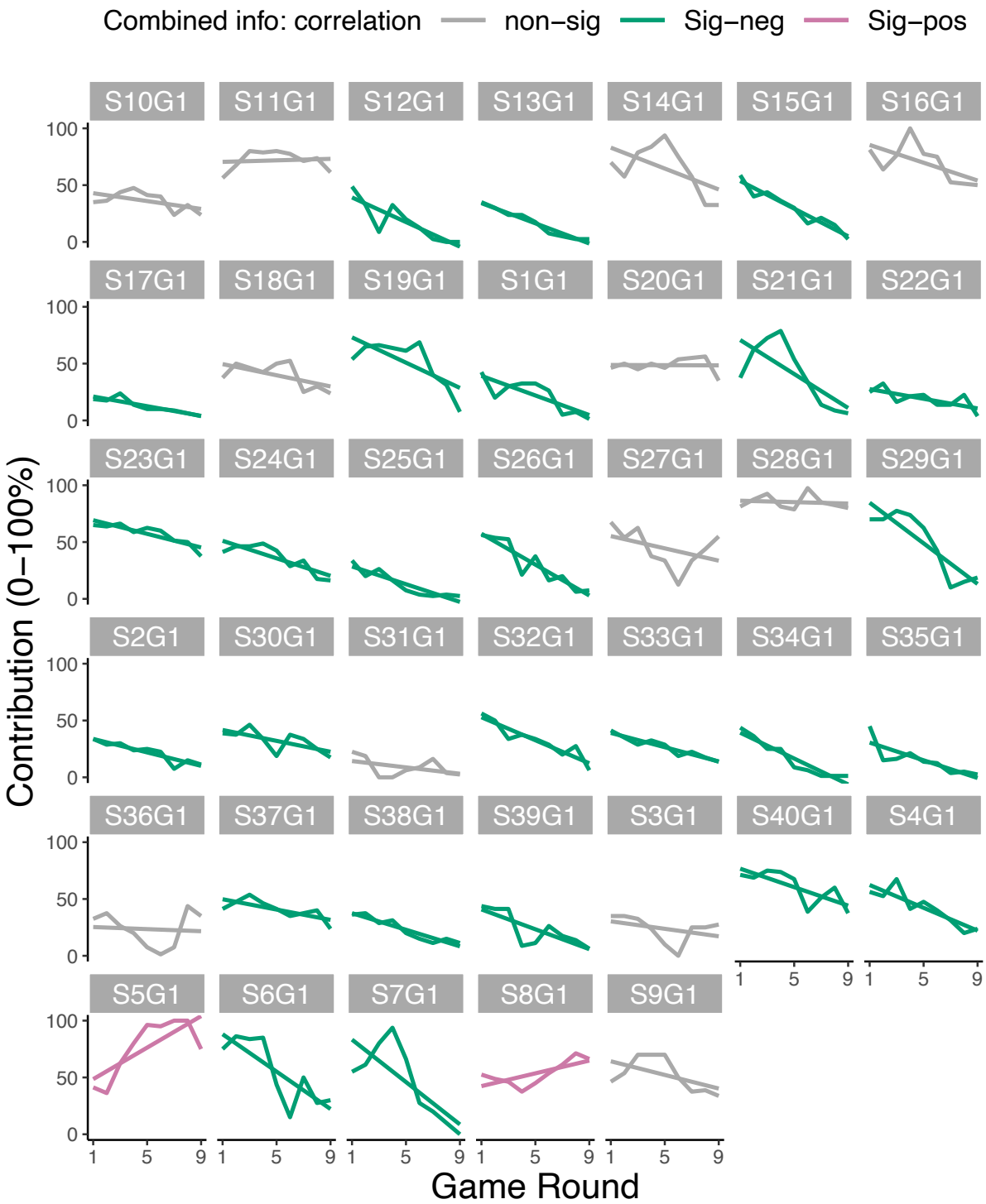
Supplementary Figure 3. Group profiles in the No-information treatment.



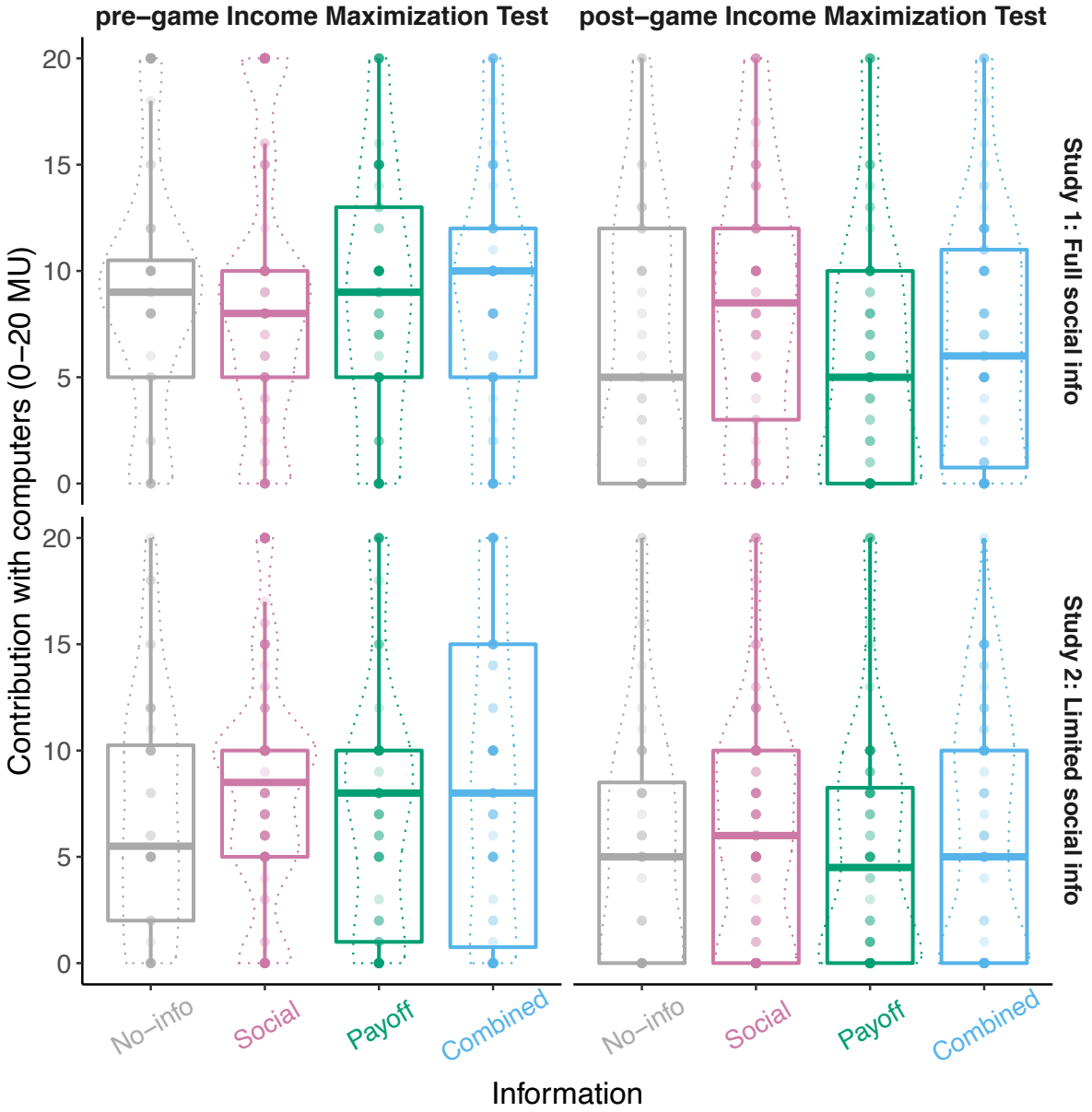
Supplementary Figure 4. Group profiles in the Social information treatment.



Supplementary Figure 5. Group profiles in the Payoff information treatment.



Supplementary Figure 6. Group profiles in the Combined information treatment.



Supplementary Figure 7. Income Maximization Tests by information treatment and study.

1041

Supplementary Table 1: Results of Study 1 with full social information.

Fixed effects	Estimate*	Std. Error	df**	t value	P value***
Reference intercept (No-info)	50.07	6.28	70.0	7.976	<.001
Social intercept	+0.32	7.69	70.0	0.042	0.967
Payoff intercept	+4.21	7.69	70.0	0.548	0.586
Combined intercept	+4.59	7.69	70.0	0.597	0.552
Reference slope (No-info)	-1.30	0.92	70.0	-1.415	0.162
Social slope	-0.04	1.12	70.0	-0.032	0.974
Payoff slope	-2.43	1.12	70.0	-2.162	0.034
Combined slope	-1.74	1.12	70.0	-1.550	0.126
Number of independent groups	70				
Number of observations	630				
Random effects					
Intercept (by group) variance	375.4	19.375			
Slope (by group) variance	6.4	2.538			
Residual	118.6	10.889			

*Response variable is the group total contribution (0-80) converted to a natural number percentage (group contribution / 0.8)

**Degrees of Freedom estimated with the Satterthwaite method

***P value is for the comparison with the reference intercept/slope

1042

Supplementary Table 2: Results of Study 2 with limited social information.

Fixed effect	Estimate*	Std. Error	df**	t value	P value***
Reference intercept (No-info)	36.04	5.40	84.02	6.675	<.001
Social intercept	+7.25	6.23	84.02	1.163	0.248
Payoff intercept	+8.39	6.72	84.02	1.248	0.216
Combined intercept	+12.90	6.72	84.02	1.919	0.058
Reference slope (No-info)	-0.86	0.67	84.00	-1.283	0.203
Social slope	-1.43	0.78	84.00	-1.844	0.069
Payoff slope	-2.60	0.84	84.00	-3.113	0.003
Combined slope	-.2.91	0.84	84.00	-3.484	<.001
Number of independent groups	84				
Number of observations	756				
Random effects					
Intercept (by group) variance	316.5	17.789			
Slope (by group) variance	3.4	1.831			
Residual	96.6	9.829			

*Response variable is the group total contribution (0-80) converted to a natural number percentage (group contribution / 0.8)

**Degrees of Freedom estimated with the Satterthwaite method

***P value is for the comparison with the reference intercept/slope

1043

1044

Supplementary Table 3: Generalized linear model analysing the second Income Maximization Test played with computers (IMT2).

Fixed effect^	Estimate [95% CI]*	df	t value	P value
Reference intercept (treatments with no payoff information)	-1.5 [-1.78, -1.29]	612	-12.327	<0.001
Treatments with payoff information	-0.27 [-0.47, -0.07]	612	-2.621	0.009
Study 2 intercept	-0.24 [-0.45, -0.04	612	-2.361	0.019
IMT1	0.11 [0.10, 0.13]	612	12.734	<0.001
Number of individuals (observations)	616 (1,232)			

^Dispersion parameter for quasibinomial family taken to be 6.277245.

^Null deviance: 5788.2 on 615 degrees of freedom.

^Residual deviance: 4582.0 on 612 degrees of freedom.

*Raw coefficients, non-exponentiated.

1045

1046

1047

1048