No. 27 may 2008

**Urs Fischbacher**
**Simon Gächter**

# Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Good Experiments

Research Paper Series
Thurgau Institute of Economics

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

# HETEROGENEOUS SOCIAL PREFERENCES AND THE DYNAMICS OF FREE RIDING IN PUBLIC GOOD EXPERIMENTS

*By* URS FISCHBACHER AND SIMON GÄCHTER*

**May 22, 2008**

**Abstract**

We provide a direct test of the role of social preferences and beliefs in voluntary cooperation and its decline. We elicit individuals' cooperation preference in one experiment and use them – as well as subjects' elicited beliefs – to make predictions about contributions to a public good played repeatedly. We find substantial heterogeneity in people's preferences. With simulation methods based on this data, we show that the decline of cooperation is driven by the fact that most people have a preference to contribute less than others. Belief formation and virtual learning do not contribute to the decline of cooperation. Universal free riding is very likely despite the fact that most people are not selfish.

Keywords: Public goods experiments, social preferences, conditional cooperation, free riding.
JEL classification: C91, C72, H41, D64.

---

* Fischbacher: University of Konstanz, P.O.Box D 131, D-78457 Konstanz, E-mail: urs.fischbacher@uni-konstanz.de and Thurgau Institute of Economics, Hauptstrasse 90, CH-8280 Kreuzlingen, E-mail: fischbacher@twi-kreuzlingen.ch. Gächter: CESifo, IZA, and University of Nottingham, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, United Kingdom. E-mail: simon.gaechter@nottingham.ac.uk. This paper is part of the MacArthur Foundation Network on Economic Environments and the Evolution of Individual Preferences and Social Norms and the Research Priority Program of the University of Zurich on "The Foundations of Human Social Behavior: Altruism versus egoism". Support from the EU-TMR Research Network ENDEAR (FMRX-CT98-0238) is gratefully acknowledged. Sybille Gübeli, Eva Poen, and Beatrice Zanella provided very able research assistance. The referee's comments and suggestions have greatly improved this paper. We also thank Nick Bardsley, Rachel Croson, Eva Eppler, Armin Falk, Ernst Fehr, Björn Frank, Sally Gschwend, Radosveta Ivanova-Stenzel, Martin Kocher, Michael Kosfeld, Andreas Kuhn, Charles Noussair, Eva Poen, Daniel Schunk, Martin Sefton, Daniel Seidmann, and various conference, workshop and seminar participants for their helpful comments. Simon Gächter thanks the University of St. Gallen, CES Munich and CMPO Bristol for the hospitality he enjoyed while working on this paper.

Numerous public goods experiments have shown that many people contribute more to the public good than pure self-interest can easily explain. However, an equally important observation is that free riding becomes paramount in repeatedly played public good experiments across various parameters (Ledyard (1995)) and also cross-culturally (Herrmann, Thöni and Gächter (2008)). The facts are clear, but their explanation is not. In this paper, we investigate the role of social preferences and beliefs for the decline of voluntary cooperation.

In the previous literature an obvious candidate for explaining the decay of cooperation was learning the free rider strategy. Andreoni (1988) showed, however, that cooperation resumed after a restart, which is inconsistent with a pure learning argument. Several papers since investigate the role of confusion and argue that contributions that are not due to confusion might possibly be explained by social preferences (e.g., Andreoni (1995); Palfrey and Prisbrey (1997); Kurzban and Houser (2005); Ferraro and Vossler (2005)). The typical approach of these experiments is indirect in the following sense: a standard public good experiment is compared to a setup in which social motives for contributing are removed. Any resulting cooperation in the latter treatment is due to confusion. The difference of contributions in the main experiment and "confused contributions" can then be possibly due to social preferences.[1]

A further indirect approach to identify the role of social preferences for the decline of cooperation is to allow subjects to revise their decisions repeatedly and to study whether the revisions depend on the contributions of the other players (Kurzban and Houser (2001); Levati and Neugebauer (2004)). Such a correlation is then interpreted as evidence for social preferences. This method has the disadvantage that there are strategic incentives and repeated game effects, which make the interpretation of correlations between other subjects'

[1] Palfrey and Prisbrey (1997) is an exception. They design an experiment that allows them estimating warm glow, altruism and errors directly. They find significant warm-glow effects and that the decay of contributions is largely due to reduced errors.

contribution and the revisions ambiguous. Finally, also experiments on "assortative regrouping" of subjects provide indirect evidence for conditional cooperation and the reasons for the decline of cooperation: In these experiments (e.g., Burlando and Guala (2005); Gächter and Thöni (2005); Page, Putterman and Unel (2005); Gunnthorsdottir, Houser and McCabe (2007)), contributors are matched up with contributors, and free riders with free riders. The results show that contributions are largely stabilized if free riders are moved out of the group. This suggests that cooperators are only *conditionally* cooperative and therefore preference heterogeneity plays an important role for explaining the decay of cooperation in randomly composed groups.

Our approach is a direct one because we measure people's cooperation preferences in a specially-designed public good game (called the "P-experiment") and observe the same people in a sequence of ten one-shot games (labeled the "C-experiment"), in which we also elicit people's beliefs about others' contributions. Such an approach is promising in our view because the fact that most people free ride eventually is a *prima facie* challenge to social preference explanations of people's cooperative behavior (see, e.g., Binmore (2006)). Moreover, our direct approach allows us to quantify how preference heterogeneity and beliefs interact in voluntary cooperation.

Our data from the P-experiment show that people differ strongly in their contribution preferences. This is consistent with previous evidence. The biggest groups of people are (i) conditional cooperators who cooperate if others cooperate, and (ii) free riders who never contribute anything, irrespective how much others contribute. We push beyond this observation of preference heterogeneity by investigating how measured preferences and beliefs are related to observed contribution behavior. We have therefore designed our experiments such that we can use the P-experiment to make a point prediction for each subject about his or her contribution in the C-experiment, given his or her beliefs.

Our conceptual separation of beliefs and preferences allows us to answer the following specific questions: First, do elicited preferences have any predictive power in explaining actual contributions? What is the role of beliefs about others' contributions for determining contributions? Second, how do people form their beliefs? Third, to what extent is the decay of cooperation determined by preference heterogeneity? The intuitive argument – suggested first by Andreoni (1995) who interpreted the decline in cooperation "to be due to frustrated attempts at kindness, rather than learning the free-riding incentives" (p. 900) – is that contributions decay because the conditional cooperators feel duped by the free riders and withdraw their contributions. Yet, previous studies (e.g., Croson (2007)) showed that contributions are also highly significantly correlated with beliefs. Thus, contributions might also decline because people lower their beliefs independently of others' contributions. This is not implausible, given evidence from other games (e.g., Weber (2003)) that even "virtual learning", that is, thinking repeatedly about the dominant strategy in our case, might make people more pessimistic about others' contributions.

Our results, which we detail in Section II, answer these questions as follows: First, contributions are significantly positively influenced by predicted contributions, that is, the elicited preferences. In addition to their preferences people's contributions also depend directly on their beliefs about others' contributions. Second, belief formation can be described as a partial adjustment of one's belief into the direction of the observed contribution of others in the previous period. There are no period effects on top of this process. As we will show, this implies that beliefs decline only if contributions decline, but not vice versa; "virtual learning" is unimportant in our data. Third, contributions decline because on average people are "imperfect conditional cooperators" who only partly match others' contributions. The presence of free rider types is not necessary for this result; contributions also decline if everyone is an imperfect conditional cooperator. However, we show that preference heterogeneity (that is, the type-composition of a group) influences the speed of the cooperative decay. In the final part of our results section we use simulation methods to understand the role of belief formation, preference heterogeneity and group composition for the decline of cooperation. Section III concludes.

## I. Design and procedures

Our basic decision situation is a standard linear public good game. The subjects are randomly assigned to groups of four people. Each participant is endowed with 20 tokens, which he or she can either keep or contribute to a "project", the public good. The payoff function is given as

$$\pi_i = 20 - g_i + 0.4\sum_{j=1}^{4} g_j \ , \qquad (1)$$

where the public good is equal to the sum of the contributions of all group members. Contributing a token to the public good yields a private marginal return of 0.4 and the social marginal benefit is 1.6. Standard assumptions therefore predict that all subjects free ride completely, that is, $g_j = 0$ for all $j$. This leads to a socially inefficient outcome.

The instructions (see Appendix) explained the public good problem to the participants. Since we want to measure subjects' preferences as accurately as possible, we also took great care to ensure that the participants understood both the rules of the game and the incentives. Therefore, after participants had read the instructions, they had to answer ten control questions. The questions tested their understanding of the comparative statics properties of (1), to ensure that participants are aware of their selfish incentives and the dilemma situation. We did not proceed until all participants had answered all questions correctly. We can thus safely assume that the participants understood the game.

Within this basic setup we conducted two types of experiments. The first type of experiment (the "P-experiment") elicits people's contribution *preferences* in a public goods game. In the second type of experiment participants make *contribution choices* in a standard linear public goods environment (labeled "C-experiment"). The C-experiment consists of ten rounds in the random matching mode. We chose a random matching protocol to minimize strategic effects from repeated play. All participants play both types of experiments. For example, participants first go through the preference elicitation experiment in the P-C sessions before making their contribution choices in the C-experiment. Our C-P sessions counterbalance the order of experiments to control for possible sequence effects. The C-P sequence allows for a particularly strong test of measured preferences because people experience ten rounds of decisions in the C-experiment before their cooperation preferences are elicited in the P-experiment.

The rationale of the P-experiment is to elicit people's cooperation preferences in the sense of eliciting their stated willingness for cooperation: To what degree are people willing to give up their free rider benefit, that is, to what extent are they prepared to cooperate given other peoples' degrees of cooperation? Being able to observe contribution preferences as a function of other group members' contribution without using deception requires observing contributions that can be contingent on others' contributions. Fischbacher, Gächter and Fehr (2001) (henceforth FGF) introduced an experimental design that accomplishes this task.[2] Since we use exactly the same method as FGF we refer the reader to FGF for all details.

The central idea of the P-experiment is to apply a variant of the so-called "strategy method" (Selten (1967)). The subjects' main task in the experiment is to indicate – *in an incentive compatible way* – how much they want to contribute to the public good *for each rounded average contribution level of other group members*. Specifically, subjects were shown a "contribution table" of the 21 possible values of the average contribution of the other group members (from 0 to 20) and were asked to state their corresponding contribution for *each* of the 21 possibilities. Since the FGF method elicits the contribution schedules in an

[2] Ockenfels (1999) developed a similar design independently of FGF.

incentive compatible way, free rider types have an incentive to enter a zero contribution for each of the 21 possible average contributions of other group members. Entering a positive contribution signals a "willingness to pay" for cooperation by foregoing the free rider benefit. We interpret this willingness to pay as a subject's cooperation preference.[3]

The experiment was only played *once*, and the participants knew this. The rationale is that we wanted to elicit subjects' preferences, without intermingling preferences with strategic considerations.

Participants in the P-C sessions (C-P sessions) were only informed *after* finishing the P-experiment (C-experiment) that they would play another experiment. When we explained the C-experiment we emphasized that the groups of four would be randomly reshuffled in each period.[4] After each period, subjects were informed about the sum of contributions in their group in that period. In addition to their contribution decisions, subjects also had to indicate their *beliefs* about the average contribution of the other three group members in the current period. In addition to their earnings from the public good experiment, we also paid subjects based on the accuracy of their estimates.[5]

[3] Our approach does not require eliciting a utility function since we do not need a complete preference order for our purposes. It is sufficient to know subjects' best replies conditional on others' contributions.

[4] The likelihood in period 1 that a player meets another player again once during the remaining nine periods was 72 percent. The likelihood that the *same* group of four players meets was 2.58 percent. However, since the experiment was conducted anonymously, subjects were unable to recognize whether they were matched with a particular player in the past.

[5] Subjects had a financial incentive for correct beliefs, but it was small to avoid hedging: If their estimation was exactly right, subjects received 3 experimental money units (≈ $0.8) in addition to their other experimental earnings. They received 2 (1) additional money units if their estimation deviated by 1 (2) point(s) from the other group members' actual average contribution, and no additional money if their estimation was off the actual contribution by more than three points.

We elicited beliefs for two reasons. First, we can assess the correlation between beliefs and contributions, which we expect to differ between types of players. Second, by evaluating an elicited schedule at the elicited belief in a given period we can make a *point prediction* about an individual's contributions in the C-experiment: if a subject in the P-experiment indicates in his or her schedule that he or she will contribute *y* tokens if the others contribute *x* tokens on average, then the prediction for this subject in the C-experiment is to contribute *y* tokens if he or she believes that others contribute *x* tokens on average.

The sequence of experiments was reversed in the C-P sessions. The comparison of results from the P-experiments in the C-P sequence with those of the P-C sequence allow us to assess the relevance of experience with the public goods game for elicited cooperation preferences.

All experiments were computerized, using the software z-Tree (Fischbacher (2007)). The experiments were conducted in the computer lab of the University of Zurich. Our participants were undergraduates from various disciplines (except economics) from the University of Zurich and the Swiss Federal Institute of Technology (ETH) in Zurich. We conducted six sessions (three in the P-C sequence and three in the C-P sequence). In five sessions we each had 24 subjects and in one 20 subjects. A post-experimental questionnaire confirmed that participants were largely unacquainted with one another. Our 140 subjects were randomly allocated to the cubicles in each session, where they took their decisions in complete anonymity from the other subjects. On average, subjects earned 35 Swiss Francs (roughly $30, including a show-up fee of 10 Swiss Francs).[6] Each session lasted roughly 90 minutes.

---

[6] During the experiment subjects earned their payoffs in "points" (according to (1) and the earnings from correct belief estimates). At the end of the experiment, we exchanged the accumulated sum of points at an exchange rate of 1 point = CHF 0.35 for the points earned in the P-experiment and at a rate of 1 point = CHF 0.07 for the points earned in the C-experiment.

## II. Results

We organize the discussion of our results as follows: In section A, we document the decline of cooperation. In section B, we present the extent of heterogeneity in people's cooperation preferences and actual contribution patterns. In the remainder of this section, we analyze behavior in the C-experiment. We show how subjects form their beliefs (section C) and how their contribution decisions are related to the elicited preferences in the P-experiment (section D). We conclude in section E with a simulation study in which we assess how the belief process and subjects' preferences affect the decline of cooperation.

### A. *The Decline of Cooperation*

Figure 1 sets the stage for our analysis, which aims to explain the decline of cooperation. The figure shows the temporal patterns of cooperation and beliefs for each of our six sessions separately. Figure 1 conveys four unambiguous messages. First, contributions and beliefs decline in six out of six sessions.[7] Second, behavior in the six sessions is very similar. Third, contributions are lower than beliefs in almost all instances. Finally, mean period contributions and beliefs are highly significantly positively correlated in all six sessions (Spearman rank correlation tests, $p<0.007$). This evidence is consistent with previous findings in public goods games (e.g., Weimann (1994); Croson (2007); Neugebauer, Perote, Schmidt and Loos (*in print*)).

[7] This finding is consistent with evidence from previous experiments in the random matching mode (e.g., Andreoni (1988); Weimann (1994); Croson (1996), Croson (2007); Burlando and Hey (1997); Sonnemans, Schram and Offerman (1999); Keser and van Winden (2000); Fehr and Gächter (2000); Park (2000)).

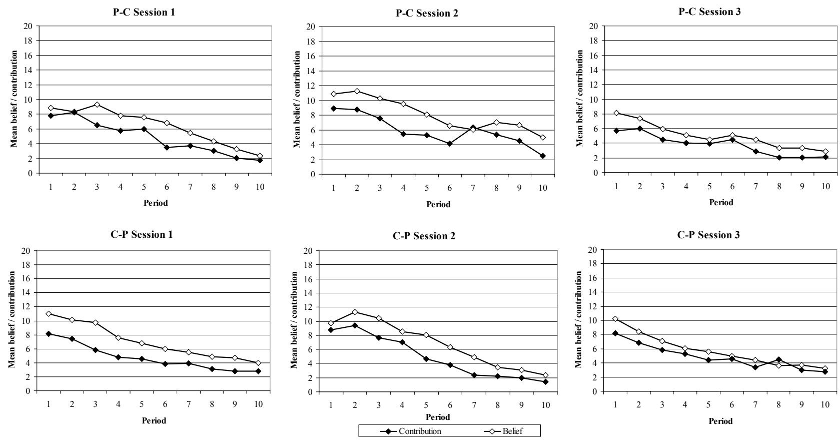



FIGURE 1. MEAN BELIEFS AND CONTRIBUTIONS OVER TIME.

In the remainder of the paper we will analyze the decline of cooperation as a result of people's beliefs as formed in the C-experiment and preferences as elicited in the P-experiment. The next step in our analysis is to look at people's cooperation preferences and to investigate the extent of heterogeneity in people's contribution behavior as a function of their beliefs in the C-experiment.

### B. *Heterogeneous Preferences and Contributions*

Recall that we have a complete contribution schedule from each subject that indicates how much he or she is prepared to contribute as a function of others' contribution. A simple way of characterizing heterogeneity is to look at the slope (of a linear regression) of the schedule and the mean contribution in the schedule. For instance, a free rider's schedule consists of zero contributions for all contribution levels of other group members. Therefore, his slope and mean contribution are zero. An unconditional cooperator, who contributes 20 tokens for all others' contribution levels, has a mean contribution of 20 and a slope of zero. A perfect conditional cooperator, who contributes exactly the amount others contribute, has a slope of one and a mean contribution of 10 tokens. Figure 2A depicts the results separately for the C-P and the P-C experiments. The *x*-axis shows the slope of the schedules and the *y*-axis the average contribution in the schedule. The dots in Figure 2A correspond to individual observations, and the size of a dot to the number of observations it represents.

Figure 2A shows two things. First, there is a large degree of heterogeneity.[8] Free riders (located at 0-0) and perfect conditional cooperators (at 1-10) are relatively the largest group of subjects. We also find a few subjects who contribute an unconditional positive amount (along the *y*-axis, at *x*=0). A large number of subjects has a positive mean contribution and a positive slope; a few subjects have a negatively-sloped schedule (that is, they contribute more the less others contribute). Second, the distribution between the C-P and the P-C sessions is very similar. Mann-Whitney tests do not allow rejecting the null hypotheses that both means and slopes are equally distributed between the treatments (p>0.87).[9] Thus, elicitation of preferences before subjects actually experienced contributions to the public good (in the P-C sessions) or after (in the C-P sessions) did not affect the elicited preferences. This is an important finding for our interpretation that the P-experiment elicits cooperation preferences. It shows that subjects in the C-P sessions who have experienced actual contribution behavior do not express different cooperation preferences than do subjects in the P-C sessions who are inexperienced in actual game playing when they express their preferences.[10]

---

[8] This evidence is consistent with other studies using different methods. See, e.g., Bardsley and Moffatt (2007), Kurzban and Houser (2005), Burlando and Guala (2005), Muller, Sefton, Steinberg and Vesterlund (forthcoming), and Herrmann and Thöni (forthcoming).

[9] In Figure 2 we looked at slope and mean contribution of a subject's schedule. However, qualitatively, we get very similar results if we look at Spearman rank order correlation coefficients, linear correlation coefficients, and slopes and intercepts of linear regressions. In all cases p-values of Mann-Whitney tests that compare the C-P and the P-C experiments yield p>0.275.

[10] The elicited contribution schedules in our study are also not significantly different from FGF ($\chi 2$-test, p=0.729). See the working paper version (Fischbacher and Gächter (2006)) for further details.

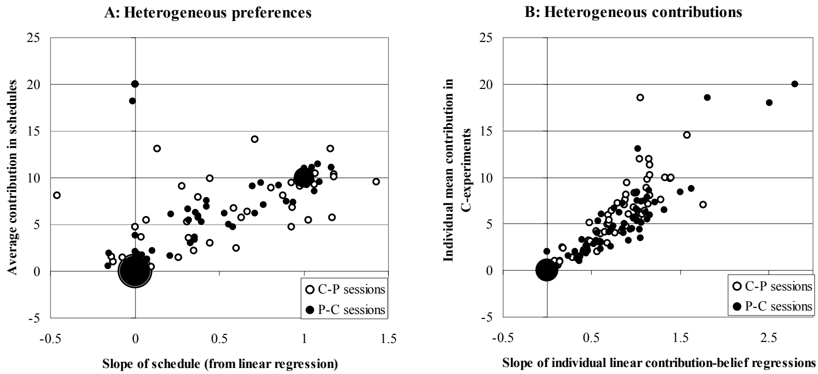



FIGURE 2. HETEROGENEOUS CONTRIBUTION PREFERENCES (PANEL A) AND ACTUAL CONTRIBUTIONS AS A FUNCTION OF BELIEFS (PANEL B).

Figure 2B shows a scatter plot of individual slopes of linear regressions of contributions on beliefs on the *x*-axis and average contributions in the C-experiment on the *y*-axis. The dot size corresponds to the number of observations. Thus, the construction of Figure 2B is similar to Figure 2A. As in Figure 2A, we distinguish between the C-P and the P-C sessions. We find no sequence effect, neither with respect to average contributions nor with respect to slopes (Mann-Whitney tests, $p > 0.21$).

Figure 2B reveals considerable heterogeneity in contribution behavior. Individual average contributions (depicted on the *y*-axis) vary between 0 and 20, although most subjects contribute less than ten tokens on average. Fourteen percent of all subjects contribute exactly zero in all ten periods. We find that the individual estimated slopes of the schedules from the P-experiment (Figure 2A) and the slopes of individual linear contribution-belief regressions in the C-experiments (Figure 2B) are highly significantly positively correlated (Spearman's $\rho=0.39$, $p=0.0000$). Average cooperation levels in the P-experiment and in the C-experiment are highly correlated as well (Spearman's $\rho=0.40$, $p=0.0000$). We interpret this as a first piece of evidence that expressed cooperation preferences and actual cooperation behavior are correlated at the individual level. Before we investigate the link between beliefs, preferences, and contributions, we look at how people form beliefs in the C-experiment.

### C. *The Formation of Beliefs*

We investigate the question how people form their beliefs about their group members' contribution in a given period with the help of six econometric models. The estimation method is OLS with robust standard errors clustered on sessions as the independent units of observation.[11]

Model 1, which only includes "Period", simply confirms the impression from Figure 1 that beliefs decline significantly over time. However, this model cannot explain *why* there is a downward trend. It is natural to assume that subjects base their belief on the other players' past contributions. Models 2 and 3 explain beliefs in a given period as a function of their co-players' contributions in the past three periods. We find that beliefs depend highly significantly on the others' contribution in the previous three periods. This fact can explain much of the decline of beliefs because the absolute value of the coefficient of the period variable is reduced from 0.753 in Model 1 to 0.243 in Model 3. A further interesting observation is that the weight of a past contribution is roughly halved from period to period.[12]

Models 2 and 3 suggest that people have a long memory when they form their beliefs. Yet, this may not be very plausible psychologically. It is more likely that subjects have access to their own previous beliefs, which they update after new information arrives. To see this, take periods 1 and 2. In period 1 a subject can only rely on his or her intuitive ("home-

---

[11] We estimated all models with random and fixed effects specifications, as well as with Tobit, with very similar results. For instance, the correlation coefficient of predicted values of the Tobit estimation and the OLS is 0.9995. Since the estimation results are very similar, we only report the OLS results for ease of interpretation.

[12] This is also true if we include contribution(t-4). We did not include more lags because our experiment only ran for ten periods.

grown") belief about others' contributions. In period 2, he or she has also an observation about others' actual contribution in period 1. A subject may therefore update his or her period 2 belief on the basis of his or her period 1 belief and the observed period 1 contributions by others. A similar logic might hold in all remaining periods.

TABLE 1—FORMING BELIEFS.

| | Dependent variable: Belief about other group members' contribution | | | | | |
|---|---|---|---|---|---|---|
| Model | 1 | 2 | 3 | 4 | 5 | 6 |
| Period | -0.753 | | -0.243 | | | -0.079 |
| | (0.061)*** | | (0.077)** | | | (0.042) |
| Others' contrib. (t-1) | | 0.476 | 0.441 | 0.439 | 0.415 | 0.394 |
| | | (0.031)*** | (0.034)*** | (0.023)*** | (0.020)*** | (0.023)*** |
| Others' contrib. (t-2) | | 0.228 | 0.190 | -0.054 | | |
| | | (0.035)*** | (0.034)*** | (0.042) | | |
| Others' contrib. (t-3) | | 0.125 | 0.084 | -0.026 | | |
| | | (0.023)*** | (0.030)** | (0.035) | | |
| Belief (t-1) | | | | 0.592 | 0.569 | 0.549 |
| | | | | (0.063)*** | (0.036)*** | (0.037)*** |
| Constant | 10.651 | 1.373 | 3.649 | 0.219 | 0.118 | 0.835 |
| | (0.661)*** | (0.327)*** | (0.909)** | (0.130) | (0.148) | (0.398)* |
| Observations | 1400 | 980 | 980 | 980 | 1260 | 1260 |
| R-squared | 0.28 | 0.37 | 0.38 | 0.62 | 0.64 | 0.64 |

OLS regressions. Robust standard errors (clustered on sessions) in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

We account for this argument in Models 4 to 6 where we include the belief a subject held in the previous period as an explanatory variable ("Belief (t-1)"). We find in Model 4 that both "Belief (t-1)" and "Others' contribution (t-1)" are highly significantly positive; more distant contributions become small and insignificant. We omit the insignificant lagged variables in Model 5 to gain more observations. Model 6 includes the period variable. The results of Models 5 and 6 are very similar to Model 4. Moreover, the sum of coefficients of "Belief (t-1)" and "Others' contribution (t-1)" is insignificantly different from 1 in Models 4 to 6 (F-tests, p>0.118). We also estimated Model 6 separately for periods 1 to 5 and periods 6 to 10. The estimated coefficients are very similar in both halves of the experiment (Chow-test, p>0.1).[13] Given these results, we can summarize the belief formation process as follows: a subject's belief in a given period is a weighted average of what he or she *believed* about others in the previous period and his or her *observation* of the others' contribution in the previous period.

The results from these regressions show that the process of belief formation is not *itself* responsible for the decline in cooperation. To isolate the process of belief formation from declining cooperation, suppose all subjects are perfect conditional cooperators whose contributions match their beliefs about others' average contribution perfectly. In this case contributions decline only if beliefs decline *per se*, that is, if the belief formation process is inherently pessimistic about others' contribution.[14] To see this, notice that for perfect conditional cooperators the average contribution is equal to their average belief and the average contribution trivially equals average others' contribution. If, furthermore, the players form their belief according to Model 5, the average belief equals AvBelief(t)= (0.569 + 0.415)* AvBelief(t-1) + 0.118. The sum of the two coefficients (0.569 and 0.415) equals 0.984, which is insignificantly different from 1 (F-test, p=0.549). This observation implies that beliefs would remain constant (or increase by 0.118 points per period). If we take the coefficient of 0.984 literally, the belief formation process *per se* can account for a decline of

[13] We also applied an Arellano-Bond linear, dynamic panel-data estimation method (Arellano and Bond (1991)). However, there is still significant second order correlation (p<0.05, Arellano-Bond test) in the residuals implying that its estimates are inconsistent (Arellano and Bond (1991), pp. 281-282). Moreover, in simulations similar to those which we discuss in Section II.E it turned out that the Arellano-Bond estimates cannot explain the data patterns at all, whereas Models 5 and 6 can.

[14] "Virtual learning" (Weber (2003)), that is learning with no feedback by just thinking about the problem for several periods is a possible reason for this "pessimism".

cooperation of at most 14 percent (i.e., $1-(1-0.984)^9=0.135$) over the nine remaining periods after period 1. Thus, beliefs decline because contributions decline and not vice versa.[15]

D. *Explaining contributions*

In this section we investigate determinants of people's contributions econometrically. We have three explanatory variables – "Period", "Predicted Contribution", and "Belief". We estimated five models which we document in Table 2. Some models differ according to the number of periods that is used in the estimations. The estimation method is OLS with robust standard errors clustered on sessions as the independent units of observation.[16]

Each of the first three models includes only one of our explanatory variables. Model 1, which only includes "Period", confirms the impression from Figure 1 that contributions decline significantly over time. Model 2 only includes the variable "Predicted contribution", which is the contribution according to the elicited schedule in the P-experiment evaluated at the subject's belief in a given period. If subjects' preferences would explain contributions perfectly, then the slope of "Predicted contribution" would be unity and the constant would equal zero. The results of Model 2 show that this is not the case. Though "Predicted contribution" is highly significantly positive its coefficient is less than unity; the constant is also positive and significant. The conclusion is that contributions are significantly influenced by subjects' preferences as elicited in the P-experiment. Preferences alone cannot account for all contributions, however.

---

[15] Neugebauer, et al. (*in print*) report experimental results that are consistent with this prediction. In one treatment subjects were informed about others' contributions in each round, and in another treatment subjects received no information about others' contributions. Their results show that contributions only decline with information but remain roughly constant in the absence of information. This finding suggests that "virtual learning" plays little role in public goods environments.

[16] As with belief formation, we estimated all models with random and fixed effects specifications, as well as with Tobit. Since the estimation results are very similar, we only report the OLS results.

Model 3 only includes beliefs as an explanatory variable. We find strong evidence for conditional cooperation: beliefs are highly significantly positively correlated with contributions.[17] The explanatory power (in terms of $R^2$) is even higher than of the variable "Predicted contribution".

Model 4 combines all variables. We find that "Predicted contribution" and "Belief" matter highly significantly. In other words, in the C-experiments, there is conditional cooperation on top of contribution preferences, but there is no significant decline in cooperation that is not explained by "Predicted contribution" and "Belief". In Models 5a to 5e we therefore concentrate on the impact of the variables "Predicted contribution" and "Belief". The models differ in the time horizon considered and in whether "confused" subjects (according to the FGF classification) are included (Model 5a) or not (Models 5b to 5e).[18]

TABLE 2—EXPLAINING CONTRIBUTIONS

| | Dependent variable: Contribution | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | 1 | 2 | 3 | 4 | 5a | 5b+ | 5c+ | 5d+ | 5e+ |
| Periods used | 1-10 | 1-10 | 1-10 | 1-10 | 1-10 | 1-10 | 1 | 1-5 | 6-10 |
| Period | -0.639 | | | -0.060 | | | | | |
| | (0.071)*** | | | (0.056) | | | | | |
| Predicted | | 0.469 | | 0.242 | 0.242 | 0.443 | 0.393 | 0.385 | 0.614 |
| contribution | | (0.069)*** | | (0.069)** | (0.069)** | (0.073)*** | (0.089)*** | (0.074)*** | (0.082)*** |
| Belief | | | 0.792 | 0.644 | 0.666 | 0.545 | 0.708 | 0.582 | 0.376 |
| | | | (0.023)*** | (0.071)*** | (0.059)*** | (0.065)*** | (0.078)*** | (0.065)*** | (0.116)** |
| Constant | 8.343 | 2.953 | -0.327 | 0.005 | -0.473 | -0.318 | -1.077 | -0.204 | -0.116 |
| | (0.545)*** | (0.144)*** | (0.215) | (0.569) | (0.244) | (0.312) | (0.509)* | (0.541) | (0.378) |
| Observations | 1400 | 1400 | 1400 | 1400 | 1400 | 1260 | 126 | 630 | 630 |
| R-squared | 0.10 | 0.16 | 0.30 | 0.34 | 0.34 | 0.38 | 0.50 | 0.33 | 0.33 |

Robust standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%.
+ Models 5b to 5e exclude (confused) subjects who, on the basis of the P-experiment could not be classified according to the FGF classification as either a "free rider", "conditional cooperator", or a "triangle contributor".

[17] This observations confirms previous findings on the importance of beliefs for contributions (e.g., Weimann (1994); Croson (2007); Neugebauer, et al. (*in print*)).

[18] We think it is justified to exclude confused subjects because their predicted contribution is not only visually unclassifiable. It also does not predict their contribution. A regression according to model 5 restricted to these types yield a coefficient of -0.078 for their predicted contribution.

Why do subjects condition their contribution decision not only on their preferences according to their predicted contribution but also take their belief into account? The P-experiment and the C-experiment, though closely related, are different games. The P-experiment is a pure one-shot game played with the strategy method, whereas the C-experiment is a sequence of ten one-shot games played in the direct response mode. Perhaps subjects see some value in keeping cooperation high, thus, cooperating for strategic reasons. Consistent with this argument we find that all types, even free riders, exhibit a positive correlation between beliefs and contributions (see Fischbacher and Gächter (2006) for details). Strategic cooperation makes only sense at the beginning. This means that the coefficient for the predicted contribution should be higher for later periods, while the coefficient for the belief should be lower. This is indeed the case, as regressions 5c to 5e show.

In the previous section, we have seen that the belief formation process *per se* cannot account for the decline in cooperation. Thus, the decline must be due to subjects' contribution behavior. Model 3 provides one answer. According to this model subjects contribute about 80 percent of what they believe the others contribute (the constant is insignificantly different from zero). According to this model, subjects would therefore contribute about $0.792^{10} \approx 9.7$ percent of the period 1 belief in period 10. When we consider Models 5a to 5e, the decline is not directly visible in these regressions, since the sum of the coefficients of "Predicted contributions" and "Belief" is close to one. However, the predicted contribution is itself dependent on the belief. Moreover, the weight subjects put on average on "Predicted contributions" is about half of the weight subjects put on "Belief". Therefore, the main source of the decline of cooperation is the fact that subjects are on average imperfect conditional cooperators, who undercut what they believe others will contribute.

E. *Simulations*

In this section we use simulation methods to understand the role of preference heterogeneity, and belief updating, for the decline of cooperation. We study variants of subjects' contribution preferences and belief update processes in order to assess which features of these processes are essential for the decline of cooperation in general, and how these features influence the speed of the decline.

All our simulations are based on a two-stage decision process. In the first stage, the simulated players form a belief about the other players' contributions. Then, players decide on a contribution, which they (partially) base on their belief. Thus, our simulation models differ in two dimensions, in how subjects form their beliefs and in how they determine their contribution. We investigate two variants of belief updating processes: "naïve belief updating" and "weighted-average belief updating". In the models with naïve belief updating, players believe that the other players contribute as much in the current period to the public good as they did in the previous period. In weighted-average belief updating people form beliefs as a weighted average of others' contributions and own beliefs in period $t$–1. We use the estimated coefficients of Model 5 in Table 1 for parameterization.

We use five different contribution models for simulating contributions.

1. In the *P-model* we assume that subjects choose their contribution according to their predicted contribution (P) only, that is, their preference schedule as elicited in the P-experiment. This model allows us to gauge the importance of mere preference interaction for the decay of cooperation.
2. In the *PB-model* the players determine their contribution according to their predicted contribution (P) as well as their beliefs (B). The weights correspond to the estimated parameters of Model 5a in Table 2. The comparison to the P-model sheds light on the importance of beliefs on top of preferences for the decay of cooperation.

3. In the *iCC-model* we assume that all players are identical conditional cooperators (iCC). That is, every simulated subject contributes according to the *average* preference schedule in the P-experiment: *Contribution* = $\alpha$ + *k***Contribution of others*. The estimates from the data of our P-experiment return $\alpha$=0.956 and *k*=0.425. Therefore, in the iCC-model *Contribution* = 0.956 + 0.425**Contribution of others*. According to this model, contributions converge to 1.66, that is, contributions above 1.66 will decline. Since the iCC-model eliminates preference heterogeneity, the comparison to the P-model – where preference heterogeneity is present – informs us therefore about the importance of preference heterogeneity for the decay of cooperation.
4. In the *iCCB-model* we assume as well that players are homogenous with respect to their preferences (players' conditional cooperation is based on the average preference schedule as in the iCC-model). However, people do not simply apply their preferences but also incorporate their beliefs directly like in the PB-model. Thus, this model informs us how for the representative player beliefs matter on top of preferences, or – in comparison with the PB model – it informs us about the role of preference heterogeneity of players when beliefs matter on top of preferences.
5. Finally, in the *pCC-model* we assume that all players are perfect conditional cooperators, that is, players match their belief exactly: *Contribution*(t) = *Belief*(t). According to this model contributions only decline if beliefs decline. This model therefore allows testing how much the belief formation process *per se* contributes to the decline of cooperation.

Since we have five basic models and two belief updating processes, we have ten models in total. In all models, we take the actual first-period contributions and beliefs as the starting values in our simulation and let contributions evolve as determined by either the schedules of our participants (in the P-model and the PB-model) or as determined by the respective decision rule of the representative agent in the iCC-model, iCCB-model and the pCC-model. For determining interactions we used the exact matching structure that was in place in the C-experiments.

Figure 3 depicts the simulation results. Panel A shows the results of our five basic models under naïve belief formation and panel B under weighted-average belief formation. We compare the simulation results to the actual average contributions over all six sessions.

First, the pCC-models predict the emergence of stable cooperation over time. The details of belief updating do not matter much for perfect conditional cooperators. These results confirm that the belief formation process *per se* does not contribute to the decline of cooperation. The comparisons to all other models show that under random matching the decay of cooperation can *only* be prevented if everyone is a perfect conditional cooperator.[19]

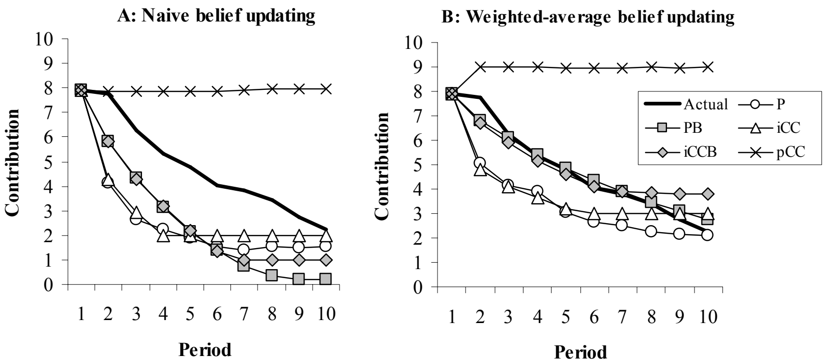



FIGURE 3. SIMULATION RESULTS.

Second, for all other models Figure 3A shows that naïve belief updating predicts a far too quick decline in cooperation. Figure 3B demonstrates that weighted-average belief updating

19 Notice, however, that it is very unlikely that four perfect conditional cooperators are matched. Among our 140 subjects we had 13 perfect conditional cooperators. The likelihood that four conditional cooperators are matched randomly is therefore less than 7.5E$^{-5}$.

does a much better job. In other words, the weight people put on their own belief in period $t$–1 relative to the weight on others' previous contributions matters significantly for the speed of the decay.

Third, in terms of closeness to the actual average contribution pattern the best model is the PB-model which combines predicted contributions and beliefs and uses weighted-average belief updating. The comparison to the P-model shows that under naïve and weighted-average belief updating, beliefs matter on top of preferences; pure preference interaction (the P-model) predicts a decline too quick compared to the actual data. As we have seen in the discussion of the regression in Table 2, beliefs could represent some strategic conditional cooperation that "slows down" the decay compared to the decay induced by preferences alone.

Fourth, we can assess the importance of preference heterogeneity in two ways, by comparing the P-model with the iCC-model and by comparing the PB-model with the iCCB-model. By construction, the iCC-model and the iCCB-model eliminate preference heterogeneity by replacing the individual preference schedules by the average preference schedule, whereas the P-model and the PB-model use the individual preference schedules. Both comparisons show that preference heterogeneity is surprisingly unimportant in explaining the decay of cooperation because the P-model and the iCC-model match each other closely; the same holds for the PB-model and the iCCB-model which are both much closer to the actual data, however. Heterogeneity matters only at the end of the experiment. In the simulations with the models with homogenous preferences contributions stop declining toward the end while the models with heterogeneous preferences correctly predict the decline also in the last periods.

We conclude this section by evaluating the predictive power of our models statistically. Table 3 collects the results. First, we calculate the correlation between the actual and the predicted contribution. The PB-model with weighted-average belief updating has the highest correlation between actual and predicted cooperation (Pearson correlation equals 0.524, and

Spearman rank correlation equals 0.460). We also ran regressions with the model prediction as regressor. Ideally, the coefficient of the model should equal 1, the constant should equal 0, and the $R^2$ should be high. This is the case for the PB-model and the iCCB-model under weighted-average belief updating (indicated by a "yes"-entry in table 3).

The second part in Table 3 reports the same statistics for the predicted beliefs. Since in all our models cooperation is based on beliefs, the models also predict beliefs. Therefore, we can apply the same econometric tests also to assess the quality of the predicted beliefs. It turns out that the PB-model with weighted-average belief updating is also the best model with respect to its predicted beliefs.

TABLE 3—MODEL PERFORMANCE SORTED BY THE SPEARMAN CORRELATION BETWEEN PREDICTED AND ACTUAL CONTRIBUTION.

| Model | | Contributions | | | | | Beliefs | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Correlations between simulated and actual contributions | | Regressions of actual contributions on simulated contributions | | | Correlations between simulated and actual beliefs | | Regressions of actual beliefs on simulated beliefs | | |
| Up-dating | Cooperation process | Spear-man | Pear-son | Slope =1 | Const =0 | $R^2$ | Spear-man | Pear-son | Slope =1 | Const =0 | $R^2$ |
| WAB | PB | 0.460 | 0.524 | yes | yes | 0.28 | 0.648 | 0.690 | yes | yes | 0.48 |
| WAB | P | 0.444 | 0.369 | | | 0.14 | 0.571 | 0.643 | | | 0.41 |
| WAB | iCC | 0.410 | 0.474 | yes | | 0.23 | 0.642 | 0.674 | yes | yes | 0.46 |
| WAB | iCCB | 0.396 | 0.500 | yes | yes | 0.25 | 0.631 | 0.678 | yes | yes | 0.46 |
| NB | iCCB | 0.378 | 0.446 | | | 0.20 | 0.595 | 0.613 | | | 0.38 |
| NB | PB | 0.361 | 0.445 | | | 0.20 | 0.577 | 0.611 | | | 0.37 |
| NB | iCC | 0.357 | 0.430 | | | 0.19 | 0.521 | 0.551 | | | 0.30 |
| NB | P | 0.356 | 0.330 | | | 0.11 | 0.366 | 0.475 | | | 0.23 |
| WAB | pCC | 0.231 | 0.378 | | | 0.14 | 0.331 | 0.491 | yes | | 0.24 |
| NB | pCC | 0.155 | 0.304 | | yes | 0.09 | 0.267 | 0.399 | | yes | 0.16 |

Note: WAB denotes "weighted-average belief updating"; NB denotes "naïve belief updating." "yes" indicates that a test fails to reject the null hypothesis that in the regressions slope=1 and constant=0, respectively.

Our simulation results suggest that the type composition of a group matters for cooperation. Groups that consist of perfect conditional cooperators should be able to maintain very high cooperation levels, whereas groups that consist mostly of free rider types or weak conditional cooperator types should exhibit low cooperation rates. These predictions are consistent with recent evidence, in which subjects are matched according to their past contributions as in Gächter and Thöni (2005), Burlando and Guala (2005) Gunnthorsdottir, et al. (2007) and Page, et al. (2005). These studies find higher contributions in the groups that were formed of subjects who contributed more in the past. Since in these high cooperator groups, cooperation was more stable than in mixed group, also average contribution of all groups was higher in treatments with segregated groups than in treatments with mixed groups.

Our results can explain these findings as follows. The regrouping mechanisms ensure that players of similar cooperativeness are matched up (either at the beginning of the experiment or various times during the experiment). Because (i) players of the respective type observe others contributing at a particular level, beliefs that others will contribute at that level are sustained, and because (ii) beliefs only decay if contributions decay, contributions actually are largely stabilized at initial levels.

Group composition effects with regard to level and decay of cooperation can occur without any assortative regroupings. Our simulation results predict this. First, the decay is slightly stronger if preferences are more heterogeneous (compare the iCC- and iCCB-models and the P- and PB-models). Second, recall that the iCC- and iCCB-model are based on *Contribution* = k**Contribution of others*, where *k* is the average slope of the contribution schedules of group members. The more free riders are in a group, the lower will this slope be. Consequently (according to Model 5c in Table 2) the lower will initial contributions be and the decay will occur faster.

The evidence is consistent with this prediction. Gunnthorsdottir, et al. (2007) observed that the cooperative decay is larger in groups with more free rider types. Similarly, Kurzban and Houser (2005) found that groups with more cooperator types exhibited higher cooperation levels than groups with more free rider types. Finally, in our own experiments, the distribution of types is not significantly different across the six sessions ($\chi^2$-test, p=0.510). It is therefore no surprise that contribution patterns are very similar in our six sessions (see Figure 1).

## III. Summary and conclusions

Our goal in this study was to provide a direct test of the role of social preferences and beliefs about others' contributions in voluntary cooperation. We achieved this by eliciting preferences in one specially-designed game (the "P-experiment") and observing contributions and beliefs in ten one-shot standard public goods games with random matching (the "C-experiments"). We found that subjects' contributions decline because on average subjects are only "imperfect conditional cooperators": they match others' contributions only partially, or, in case of free riders, not at all. We have shown that the belief formation process contributes only little – if anything at all – to the decline of cooperation. Thus, the decline of cooperation can only be prevented if average conditional cooperation is sufficiently high. Since there are almost no subjects who contribute more than the others, this means that stable cooperation can only emerge if all players are perfect conditional cooperators.

Our results reinforce the theoretical prediction by recent models of social preferences that in the absence of punishment opportunities, voluntary cooperation will end in almost universal free riding behavior.[20] This holds despite the fact that a majority of people is *not* motivated by selfishness. We think this result is important in view of arguments that deviations from income-maximizing Nash-equilibria are largely due to inexperience and lack of trial-and-error learning. Binmore (2006) makes this viewpoint very clear in a recent comment on social preference explanations of experimental results:

[20] Consistent with this prediction, Herrmann, et al. (2008) found in public goods experiments they ran in sixteen subject pools around the world that cooperation collapsed in all but one subject pools.

> "… There is a huge literature which shows that adequately rewarded laboratory subjects learn to play income-maximizing Nash equilibria in a wide variety of games – provided they have gained sufficient experience of the game and the way that other subjects play".

After Binmore has cited the standard result of eventual free riding in repeated public goods experiments in support of the above argument he argues:

> "I emphasize the standard results in Public Goods games because the orthodox view among mainstream economists and game theorists (…) is not that the learning or trial-and-error adjustment that might take place during repeated play (against a new opponent each time) is a secondary phenomenon to which conclusions may or may not be sensitive. On the contrary, the fact that laboratory subjects commonly adapt their behavior to the game they are playing as they gain experience is entirely central to our position."

Our results from the P-experiment and the C-experiment, which separate preferences, beliefs and behavior, uncover how experience in the public goods game matters. First, beliefs only decline if contributions decline; "virtual learning" (thinking about the dominant strategy in this case) is unimportant. Second, contributions are determined significantly by people's contribution preferences and beliefs. This is also true for the initial contributions, where people have not yet made any experience. If people were confused and did not know what to do, such a correlation would be unlikely. Third, the preference-type composition of groups determines the contribution level and the contribution path systematically. In groups with more free rider types the decay happens faster than in groups with fewer free riders. In groups that consist of cooperators only, contributions are stabilized at high levels. These are very unlikely outcomes if contributions were solely due to trial-and-error learning. Fourth, contributions decline (in randomly composed groups) because the majority of people are imperfect conditional cooperators and because the conditional cooperators have no punishment opportunity other than withdrawing their own contribution if they are duped by free riders. Thus, after some time all types *behave* like income-maximizing free riders, even though only a minority is motivated by income-maximization alone.

**REFERENCES**

**Andreoni, James.** 1988. "Why Free Ride - Strategies and Learning in Public-Goods Experiments." *Journal of Public Economics*, 37(3): 291-304.

**____.** 1995. "Cooperation in Public-Goods Experiments - Kindness or Confusion?" *American Economic Review*, 85(4): 891-904.

**Arellano, Manuel and Bond, Stephen R.** 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2): 277-97.

**Bardsley, Nicholas and Moffatt, Peter G.** 2007. "The Experimetrics of Public Goods: Inferring Motivations from Contributions." *Theory and Decision*, 62(2): 161-93.

**Binmore, Ken.** 2006. "Economic Man - or Straw Man?" *Behavioral and Brain Sciences*, 28(6): 817-18.

**Burlando, Roberto M. and Guala, Francesco.** 2005. "Heterogeneous Agents in Public Goods Experiments." *Experimental Economics*, 8(1): 35-54.

**Burlando, Roberto M. and Hey, John D.** 1997. "Do Anglo-Saxons Free Ride More? ." *Journal of Public Economics*, 64: 41-60.

**Croson, Rachel.** 2007. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games." *Economic Inquiry*, 45(2): 199-216.

**Croson, Rachel** 1996. "Partners and Strangers Revisited." *Economics Letters*, 53(1): 25-32.

**Fehr, Ernst and Gächter, Simon.** 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*, 90(4): 980-94.

**Ferraro, Paul J. and Vossler, Christian, A.** 2005. "The Dynamics of Other-Regarding Behavior and Confusion: What's Really Going on in Voluntary Contributions Mechanism Experiments?" Experimental Laboratory Working Paper Series #2005-001, Dept. of Economics, Andrew Young School of Policy Studies, Georgia State University.

**Fischbacher, Urs.** 2007. "Z-Tree: Zurich Toolbox for Readymade Economic Experiments." *Experimental Economics*, 10: 171-78.

**Fischbacher, Urs and Gächter, Simon.** 2006. "Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods." CeDEx Discussion Paper 2006-01, University of Nottingham.

**Fischbacher, Urs; Gächter, Simon and Fehr, Ernst.** 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71(3): 397-404.

**Gächter, Simon and Thöni, Christian.** 2005. "Social Learning and Voluntary Cooperation among Like-Minded People." *Journal of the European Economic Association*, 3(2-3): 303-14.

**Gunnthorsdottir, Anna; Houser, Daniel and McCabe, Kevin.** 2007. "Disposition, History and Contributions in Public Goods Experiments." *Journal of Behavior and Organization*, 62: 304-15.

**Herrmann, Benedikt and Thöni, Christian.** forthcoming. "Measuring Conditional Cooperation: A Replication Study in Russia." *Experimental Economics*.

**Herrmann, Benedikt; Thöni, Christian and Gächter, Simon.** 2008. "Antisocial Punishment across Societies." *Science*, 319: 1362-67.

**Keser, Claudia and van Winden, Frans.** 2000. "Conditional Cooperation and Voluntary Contributions to Public Goods." *Scandinavian Journal of Economics*, 102(1): 23-39.

**Kurzban, Robert and Houser, Daniel.** 2001. "Individual Differences in Cooperation in a Circular Public Goods Game." *European Journal of Personality*, 15: S37-S52.

**____.** 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(5): 1803-07.

**Ledyard, John O.** 1995. "Public Goods: A Survey of Experimental Research." In *The Handbook of Experimental Economics,* Alvin E. Roth and John H. Kagel, 111-81. Princeton: Princeton University Press.

**Levati, M. Vittoria and Neugebauer, Tibor.** 2004. "An Application of the English Clock Market Mechanism to Public Goods Games." *Experimental Economics*, 7(2): 153-69.

**Muller, Laurent; Sefton, Martin; Steinberg, Richard and Vesterlund, Lise.** forthcoming. "Strategic Behavior and Learning in Repeated Voluntary-Contribution Experiments." *Journal of Economic Behavior & Organization*.

**Neugebauer, Tibor; Perote, Javier; Schmidt, Ulrich and Loos, Malte.** *in print*. "Self-Biased Conditional Cooperation: On the Decline of Cooperation in Repeated Public Goods Experiments." *Journal of Economic Psychology*.

**Ockenfels, Axel.** 1999. *Fairness, Reziprozität Und Eigennutz - Ökonomische Theorie Und Experimentelle Evidenz*. Tübingen: Mohr Siebeck.

**Page, Talbot; Putterman, Louis and Unel, Bulent.** 2005. "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency." *Economic Journal*, 115: 1032-52.

**Palfrey, Thomas R. and Prisbrey, Jeffrey E.** 1997. "Anomalous Behavior in Public Goods Experiments: How Much and Why?" *American Economic Review*, 87(5): 829-46.

**Park, Eun-Soo.** 2000. "Warm-Glow Versus Cold-Prickle: A Further Experimental Study of Framing Effects on Free-Riding." *Journal of Economic Behavior and Organization*, 43: 405-21.

**Selten, Reinhard.** 1967. "Die Strategiemethode Zur Erforschung Des Eingeschränkt Rationalen Verhaltens Im Rahmen Eines Oligopolexperimentes." In *Beiträge Zur Experimentellen Wirtschaftsforschung,* Heinz Sauermann, 136-68. Tübingen: J.C.B. Mohr (Paul Siebeck).

**Sonnemans, Joep; Schram, Arthur and Offerman, Theo.** 1999. "Strategic Behavior in Public Good Games: When Partners Drift Apart." *Economics Letters*, 62: 35-41.

**Weber, Roberto A.** 2003. "'Learning' with No Feedback in a Competitive Guessing Game." *Games and Economic Behavior*, 44: 134-44.

**Weimann, Joachim.** 1994. "Individual Behavior in a Free Riding Experiment." *Journal of Public Economics*, 54: 185-200.

# Appendix: Instructions for the experiment

*This is a translation of the original German version. We present the instructions of the P-C experiments here; those of the C-P experiments were adapted accordingly. They are available upon request.*

## Instructions for the P-Experiment

You are now taking part in an economics experiment financed by the Swiss Science Foundation. If you read the following instructions carefully, you can – depending on your decisions – earn some more money in addition to the 10 Francs, which you can keep in any case. The entire amount of money which you earned with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions 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 the 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.

We will not speak of Francs during the experiment, but rather of points. Your whole income will first be calculated in points. At the end of the experiment, the total amount of points you earned will be converted to Francs at the following rate:

**1 point = 35 centimes.**

All participants will be divided in groups of four members. **Except for us - the experimenters - no one knows who is in which group.**

We describe the exact experiment process below.

### 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 description of the decision situation that help you to understand the decision situation.

You will be a member of a group consisting of **4 people**. Each group member has to decide on the allocation of 20 points. You can put these 20 points 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 point for each point you put into your private account.** For example, if you put 20 points into your private account (and therefore do not invest into the project) your income will amount to exactly 20 points out of your private account. If you put 6 points into your private account, your income from this account will be 6 points. **No one except you earns something from your private account.**

### Your income from the project

**Each group member will profit equally from the amount you invest into the project.** On the other hand, you will also get a payoff 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 × 0.4* |

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

### 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)*<br>*+ Income from the project (= 0.4 × sum of all contributions to the project)* |
| *Total income* |

**Control questions:**

Please answer the following control questions. They will help you to gain an understanding of the calculation of your income, which varies with your decision about how you distribute your 20 points. *Please answer all the questions and write down your calculations.*

1. Each group member has 20 points. Assume that none of the four group members (including you) contributes anything to the project.
   What will *your* total income be? ___________
   What will the total income of the *other* group members be? ___________

2. Each group member has 20 points. You invest 20 points in the project. Each of the other three members of the group also contributes 20 points to the project.
   What will *your* total income be? ___________
   What will the total income of the *other* group members be? ___________

3. Each group member has 20 points. The other 3 members contribute a total of 30 points to the project.

   a) What will *your* total income be, if you – in addition to the 30 points – invest 0 points into the project?
   ***Your* Income** ___________

   b) What will *your* total income be, if you – in addition to the 30 points – invest 8 points into the project?
   ***Your* Income** ___________

   c) What will *your* total income be, if you – in addition to the 30 points – invest 15 points into the project?
   ***Your* Income** ___________

4. Each group member has 20 points at his or her disposal. Assume that you invest 8 points to the project.

   a) What is your total income if the other group members – in addition to your 8 points – contribute another 7 points to the project?
   ***Your* Income** ___________

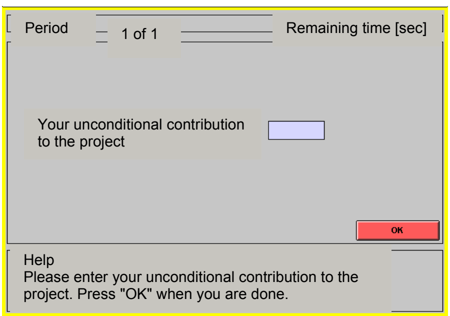
   b) What is your total income if the other group members – in addition to your 8 points – contribute another 12 points to the project?
   ***Your* Income** ___________

   c) What is your income if the other group members – in addition to your 8 points – contribute another 22 points to the project?
   ***Your* Income** ___________

**The Experiment**

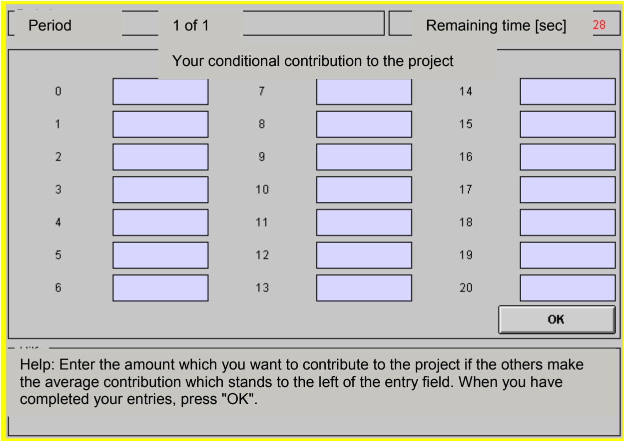
The experiment includes the decision situation just described to you. You will be paid at the end of the experiment based on the decisions you make in this experiment. The experiment will only be conducted **<u>once</u>**. As you know, you will have 20 points at your disposal. You can put them into a private account or you can invest them into a project. Each subject has to make **two types** of decisions in this experiment, which we will refer to below as the **"unconditional contribution"** and **"contribution table"**.

- You decide how many of the 20 points you want to invest into the project in the **unconditional** contribution. Please indicate your contribution in the following computer screen:




After you have determined your unconditional contribution, please click "OK".

- Your second task is to fill in a "contribution table" where you indicate how many tokens **<u>you</u> want to contribute** to the project **for each possible average contribution of the <u>other</u> group members** (rounded to the next integer). You can condition your contribution on that of the other group members. This will be immediately clear to you if you take a look at the following table. This table will be presented to you in the experiment:




The numbers are the possible (rounded) average contributions of the **other** group members to the project. You simply have to insert how many tokens you will contribute to the project into each input box – conditional on the indicated average contribution. **You have to make an entry into each input box**. For example, you will have to indicate how much you contribute to the project if the others contribute 0 tokens to the project, how much you contribute if the others contribute 1, 2, or 3 tokens, etc. You can insert **any integer numbers from 0 to 20** in each input box. Once you have made an entry in each input box, click "OK".

After all participants of the experiment have made an unconditional contribution and have filled in their contribution table, a random mechanism will select a group member from every group. Only **the contribution table will be the payoff-relevant decision for the randomly determined subject**. Only the **unconditional contribution** will be the payoff-relevant decision for the **other three group members** not selected by the random mechanism. You obviously do not know whether the random mechanism will select you when you make your unconditional contribution and when you fill in the contribution table. You will therefore have to think carefully about both types of decisions because both can become relevant for you. Two examples should make this clear.

**<u>EXAMPLE 1</u>**: Assume that **the random mechanism selects you. This implies that your relevant decision will be your contribution table**. The unconditional contribution is the relevant decision for the other three group members. Assume they made unconditional contributions of 0, 2, and 4 tokens. The average contribution of these three group members, therefore, is 2 tokens. If you indicated in your contribution table that you will contribute 1 token if the others contribute 2 tokens on average, then the total contribution to the project is given by 0+2+4+1=7 tokens. All group members, therefore, earn 0.4×7=2.8 points from the project plus their respective income from the private account. If, instead, you indicated in your contribution table that you would contribute 19 tokens if the others contribute two tokens on average, then the total contribution of the group to the project is given by 0+2+4+19=25. All group members therefore earn 0.4×25=10 points from the project plus their respective income from the private account.

**<u>EXAMPLE 2:</u>** Assume **that the random mechanism did not select you,** implying that **the unconditional contribution is taken as the payoff-relevant decision** for you and two other group members. Assume your unconditional contribution is 16 tokens and those of the other two group members are 18 and 20 tokens. Your average unconditional contribution and that of the two other group members, therefore, is 18 tokens. If the group member whom the random mechanism selected indicates in her contribution table that she will contribute 1 token if the other three group members contribute on average 18 tokens, then the total contribution of the group to the project is given by 16+18+20+1=55 tokens. All group members will therefore earn 0.4×55=22 points from the project plus their respective income from the private account. If, instead, the randomly selected group member indicates in her contribution table that she contributes 19 if the others contribute on average 18 tokens, then the total contribution of that group to the project is 16+18+20+19=73 tokens. All group members will therefore earn 0.4×73=29.2 points from the project plus their respective income from the private account.

**The random selection of the participants** will be implemented as follows. Each group member is assigned a number between 1 and 4. As you remember, a participant, namely the one with the number 11, was randomly selected at the very beginning of the experiment. This participant will throw a 4-sided die **after** all participants have made their unconditional contribution and have filled out their contribution table. The resulting number will be entered into the computer. If participant 11 throws the membership number that was assigned to you, then your contribution table will be relevant for you and the unconditional contribution will be the payoff-relevant decision for the other group members. Otherwise, your unconditional contribution is the relevant decision.

---

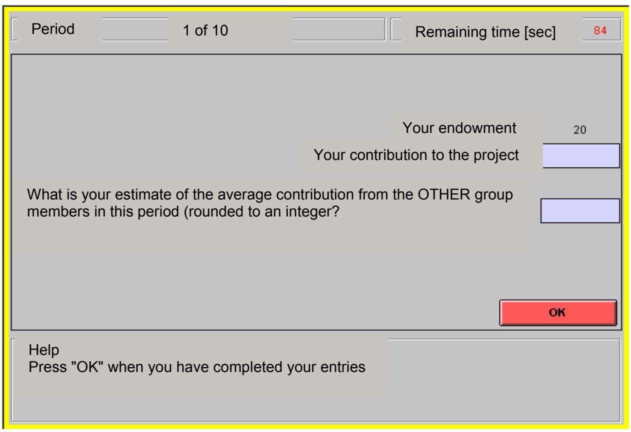
## Instructions for the C-Experiment

We will now conduct another experiment. This experiment lasts **10 periods**, in which you and the other group members have to make decisions. As in the other experiment, every group consists of **4 people**. The formation of the group changes at random after every period. **So your group consists of <u>different</u> people in all 10 periods.** The whole experiment is finished after these 10 periods,.

The decision situation is the same as that described on page 2 of the instructions of the previous experiment. Each member of the group has to decide about the usage of the 20 points. You can put these 20 points into your private account or you can invest them fully or partially into a project. Each point you do not invest into the project is automatically placed into your private account. Your income will be determined in the same way as before. Reminder:

*Income from your private account* (= *20 – contribution to the project*)
+ *Income from the project* (= *0.4 × sum of all contributions to the project*)
*Total income*
**1 point = 7 centimes!**

The decision screen, which you will see in every period, looks like this:




As you can see, you have to make two inputs:

1. First you have to **decide on your contribution to the project**, that is, you have to decide how many of the 20 points you want to contribute to the project, and how many points you want to put into your private account. This decision is the same as the unconditional contribution of the previous experiment. You only make unconditional decisions in this experiment. There is **no contribution table**.

2. Afterwards you have to estimate the average contribution to the project (rounded to an integer) of the other three group members of this period. You will be paid for the accuracy of your estimate:

- If your estimate is exactly right (that is, if your estimate is **exactly** the same as the actual average contribution of the other group members), you will get **3 points** in addition to your other income from the experiment.
- If your estimate deviates by one point from the correct result, you will get 2 additional points.
- A deviation by 2 points still earns you 1 additional point.
- If your estimate deviates by 3 or more points from the correct result, you will not get any additional points.

After these 10 periods are over, the whole experiment is finished and you will receive:

+ your income from the first experiment
+ your income from the second experiment (including your income from your correct estimates)
= total income from both experiments

+ 10 Francs show up fee !

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
**www.twi-kreuzlingen.ch**