Overcontribution and decay in public goods experiments: a test of the heterogeneous agents hypothesis*

Roberto M. Burlando
Department of Economics, University of Torino
and School of Psychology, University of Exeter
roberto.burlando@unito.it

Francesco Guala
Cognitive Science Laboratory, University of Trento
and Centre for Philosophy of the Social Sciences, University of Exeter
f.guala@ex.ac.uk

Last revised: 27 September 2002

Abstract: standard attempts to explain the phenomenon of decaying contribution in repeated linear public goods games are based on a 'representative agent' approach, with either selfish or altruist agents and an 'error' component. In this paper we try to test by purely experimental means the alternative hypothesis that in experimental public goods games there are at least three types of player: free riders, cooperators, and reciprocators. We try to identify the various types by means of four classification methods, and then play the public goods game with homogeneous groups. We observe that (1) the average contribution level is enhanced in this setting; (2) the decay phenomenon is replicated in groups of 'pure' free riders, whereas in groups of cooperative and reciprocating players the contribution is high and fairly stable throughout the game.

JEL Classification: H41, C92.

* The experiments reported in this paper were financed by the Cognitive Science Laboratory and the Computational and Experimental Economics Laboratory of the University of Trento. We would like to thank Marco Tecilla, Macrina Marchesin and Dino Parisi for their help in running the experiments, and Luigi Mittone, director of CEEL, for letting us use the laboratory's resources.

1. Introduction

'Overcontribution' in linear Public Goods (PG) experiments is by now a well-established phenomenon. It also established that the level of contribution tends to diminish with repetition. This phenomenon of declining contribution is sometimes referred to as 'decay'. However, overcontribution does not disappear completely, even after up to 60 rounds (cf. Ledyard 1995).

Standard theory assumes perfectly rational self-interested individuals, and rules out any contribution to the public good. In order to explain the anomaly, some scholars have focused on *altruism* as the main explanation of overcontribution. Altruism, however, leaves the decay phenomenon unexplained. Some recent work tries to combine the 'altruism' explanation with an *error and learning* hypothesis. Palfrey and Prisbey (1996, 1997) and Anderson et al. (1998), for example, modify the standard economic model by means of an 'altruism' parameter together with a stochastic component. The decay of contribution is thus interpreted as a process of error-elimination, or 'discovery' of the structure of the game, or of one's 'true' preferences.[1]

A distinctive characteristic of such models is their neglect of any interaction between players: altruism and learning are supposed to work the same way for each subject regardless of the behaviour of the other players in the PG game. Another approach, based on a plausible psychological hypothesis, is that the decision to contribute is affected by the context of the game, and in particular by the behaviour of the other members of the experimental group. 'Reciprocating players' contribute if and only if (some of) the others do the same. They may, in particular, link their level of contribution to the average contribution of the group – hence the well-known phenomenon of 'splitting' (contributing a sum between Nash and the social optimum). Part of the decay phenomenon, then, could be due to the presence of reciprocators who are prompted by the context, e.g. the presence of some free riders, to slide towards self-interested behaviour.

[1] This approach has the advantage of promising a unified account of a number of 'similar' experimental phenomena, such as convergence to efficient equilibria in double-oral auctions and reduction of intransitivities in repeated preference reversal experiments (cf. Plott 1995).

Despite several attempts in this direction, reciprocating agents turn out to be remarkably difficult to model.[2] Their existence, and their relevance for the decay of overcontribution, can however be established in a non-theoretical way, by purely experimental means. Suppose that – contrary to standard economic theory and to the models of altruism plus error cited above – the decay phenomenon resulted from the combination of different types of players. Suppose the experimental sample were composed by 'pure' free riders and 'pure' cooperators, as well as by reciprocators willing to give conditional on the others' contribution. An important implication of this scenario (the 'heterogeneous agents hypothesis'[3]) is that it makes decay dependent on the composition of the experimental groups. If there are enough free riders in the experimental population, uniformly distributed across the groups, the decay phenomenon gets triggered. But if we could somehow isolate free riders from players with a cooperative or reciprocating attitude, then we should observe a quick decay towards Nash in groups of free riders, vs. a more stable contribution rate in groups made up entirely of reciprocators and/or 'pure' cooperators.

In this paper we report an experiment aimed at testing the heterogeneous agents hypothesis. We implement a within-subjects design where the same individuals participate first in a repeated linear PG experiment with groups of *heterogeneous* subjects, and then in a repeated linear PG experiment with *homogeneous* groups. The hypothesis implies that in the second experiment contribution will be much lower in groups of free riders than in groups of ('pure' or 'conditional') cooperators, and that reciprocators will mutually support their level of contribution. The higher rate of contribution in groups of cooperators and in groups of reciprocators will at least partly compensate for the lower rate of contribution in groups of free-riders. Unless free riders constitute the majority of the population, therefore, we should observe a higher overall rate of contribution in the second than in the first experiment.

The most difficult task in running such an experiment is to identify and classify subjects according to 'types'. In this experiment we have combined four sources of evidence: the

---

[2] Cf. Sugden (1984), Rabin (1993), Levine (1998); see also Fehr and Fischbacher (2002) for a general discussion.

[3] This hypothesis has a long history in economic and social psychology (see e.g. Johnson and Nohrem-Hebeisem 1979, Webley et al. 1988) but has entered the economics literature only recently; its implications and relevance has been pointed out with respect to various policy issues by Hekcman (2001a, 2001b). Recent economic experimental evidence on PG games pointing in this direction includes Andreoni (1995b), Burlando and Webley (1999), Offerman et al. (1996), Weiman (1994),.

‘Strategy Method’ used by Fischbacher et al. (2001), the ‘Decomposed Game Technique’ used by Offerman et al. (1996), various measures of behaviour in a repeated linear PG game (along the lines of Burlando & Webley, 1999), and a questionnaire. Section 2 is devoted to illustrating these techniques and the experimental design in general. Section 3 contains the results of the experiment, and section 4 concludes.

2. Experimental design

The experiment was run at the University of Trento (Italy) in May 2002, and involved 92 subjects (mostly, but not entirely, undergraduates from the School of Economics and Business). The subjects were recruited by means of flyers, and were asked to come to the laboratory twice, with exactly a one-week interval between the first and the second session.[4] Overall subjects spent about 80 minutes in the lab (about 50 for the first session and 30 for the second one) and earned on average 20.50 Euros[5]. The experiment was run entirely by computerised means, with up to 20 subjects sitting in the same room, in front of terminals isolated by means of partitions.[6] The experimental currency was expressed in ‘tokens’, with 1 token = 1 cent of Euro. The first session consisted of four different tasks. The subjects were allocated randomly to terminals, and provided with a sheet of instructions for the first task. Each subject was instructed to press a key when she had finished reading the instructions; when all the subjects had done so, the experimenter asked if anybody wanted to ask any question. Then, the first experiment began. (The same procedure was followed before each experimental task.)

2.1 Session 1, task 1: Strategy Method

The so-called ‘Strategy Method’ was first used by Fischbacher et al. (2001) in an attempt to observe the phenomenon of reciprocation (or ‘conditional cooperation’) directly. The environment for this task is a linear PG game with payoff function

$$\pi_i = 200 - g_i + 0.5\sum_{j=1}^{4} g_j \ , \qquad (1)$$

[4] Subjects were told in advance that payment was strictly conditional on participating to both sessions, and that earnings from both sessions would be paid immediately after the second one. Those who didn’t show up at the second session lost their earnings from the first session, and their data were disregarded.
[5] This is more than what an average Italian student can earn in a part-time job.
[6] The software for the experiment was created by Marco Tecilla at CEEL.

where 200 is the total number of tokens to be shared between a ‘private’ (200 – *g*) and a ‘public’ account (*g*).[7] Once subjects have been made familiar with the situation, they are asked to take two types of decision: first, they are asked to make an ‘unconditional contribution’, i.e. to decide how much they would like to contribute in a standard one-shot PG game where each player, at the moment of taking her decision, doesn’t know how much the other players have contributed. Secondly, subjects are asked to fill in a (conditional) ‘contribution table’, i.e. to indicate how much they *would* be willing to contribute *if* they knew that the other members of their group had, on average, contributed to the public good a given amount. This question is iterated, varying the amount hypothetically contributed by the ‘others’ (in discrete intervals) from 0 to 200. In other words, subjects are asked to make a series of ‘conditional contributions’ in addition to the unconditional one just indicated. Clearly the data from the ‘contribution table’ are particularly useful in order to identify reciprocating players. Participants know that after the decisions have been made subjects will be randomly allocated to groups of 4 players, one of which will be selected at random as the one who will actually play the conditional contribution task, based on the other three players' unconditional decisions. The actual rewards are then calculated (and communicated to the players) according to the payoff function above. This way, both decisions (conditional and unconditional) are made relevant for the final result, and monetary incentives are provided for all members of the group.[8]

2.2 Session 1, task 2: Decomposed Game

The Decomposed Game technique has been widely used by psychologists and (more recently) economists in order to measure attitudes towards cooperation. Subjects are asked to make 24 choices between pairs of allocations. Each subject knows that she has been paired with another participant who will remain the same, but unknown, throughout the game. Each allocation consists of a number of tokens paid to yourself and another sum paid to the other player. The token amounts can be positive or negative. A typical choice may involve, e.g., a combination A = (75, -130) vs. B = (39, -145), where one

---

[7] In this phase we used (with minor adaptations) instructions and control questions provided by Urs Fischbacher, whom we would like to thank. With respect to the Fischbacher et al. experiment we raised the marginal payoff function of contributions to the public goods from 0.4 to 0.5 tokens. After the unconditional contribution task, we also asked subjects ‘how much do you think the other members of the group have contributed to the public good, on average?’

[8] See Fischbaher et al. (2001) for a more detailed analysis of this procedure. Unlike in the original one, in our experiment the random selection mechanism was run by the computer itself.

must choose between gaining either 75 or 39 tokens, with related losses on the other's part of either 130 or 145 tokens. The total sum ('own' plus 'other') allocated is not constant over the 24 combinations. (The full list of A, B combinations is standard and can be found in Appendix 1.) There is no feedback concerning the other's choices, until all participants have finished this task. The final payoff is obtained by combining the 24 choices of each subject with those of the other player.[9]

At the end of the experiment we take the 24 vectors chosen by each subject and add them up to obtain his or her own 'motivational vector'. We use the standard classification criteria used in the previous literature:[10] the motivational vector is placed on the standard 'value orientation circle' (see Appendix 1) and classified accordingly. The length of the motivational vector is used as a measure of a subject's consistency. A perfectly consistent subject (i.e. a subject who always chooses the alternative lying closest to her own motivational vector) should have a vector length of 300. Subjects with a vector length of 150 or less (less than 50% of the maximal length) thus display a considerable degree of inconsistency or confusion.

2.3. Session 1, task 3: Repeated linear PG experiment (heterogeneous groups)

The third task is a repeated linear PG experiment, with the same payoff function as in the first task (the Strategy Method), except that the individual endowment at each round is 20 tokens instead of 200. A new set of instructions reminds subjects about the environment, and states clearly that the groups (of 4 players) will be different from those of tasks one and two. The PG game will be played for 23 rounds, with the first three rounds for training (the payoffs do not count) and 20 rounds for real. After each round the subjects are given feedback about their total earnings, their earnings in the previous round, and the average contribution level of their group. The number of rounds that have already been played is displayed at the bottom of the screen.

---

[9] For example: suppose subject *i* accumulates a total payoff of 500 ('own') and allocates to the 'other' a total payoff of 100; his counterpart *j*, in contrast, accumulates a total payoff of 550 ('own') and allocates to the other a total payoff of –40. In the end, *i* gets 500–40=460, whereas *j* receives 550+100=650 tokens.

[10] Cf. Griesinger and Livingston (1973), Liebrand (1984), Offerman et al. (1996).

2.4. Session 1, task 4: Questionnaire

Before leaving the room, subjects were required to fill in a questionnaire (on their PCs). The first four questions were open, with enough space to write an articulate answer: (1) "What were you trying to do in the experiment (in other words: what were your goals of objectives)?" (2) "Did you achieve your objectives?" (3) "What were the other members of your group trying to do (what were their objectives)?" (4) "What was the scope of this experiment (in other words, what were the experimenters trying to discover)?"

Then, each subject was asked to indicate her level of agreement with four statements, on a 1-7 scale: (5) "This experiment requires a great concentration effort"; (6) "The rules of the game were explained clearly and were understandable"; (7) "In this experiment one must try to work together with others, in order to have everyone end up with more money"; (8) "In this experiment, everyone's earnings depend on the decisions of all members of the group". The relevant questions for us were (1), (2), (3), and (7). Question (7) was used by Brandts and Schram (2001) as an indicator of individual attitude towards cooperation; by means of the first three questions (taken from Burlando and Webley, 1999) we hoped to collect useful data about the strategies implemented during the first three tasks of the experiment.

2.5. Classification criteria

We used data from the four tasks described above in order to classify individual players. For the purposes of our research, we needed to divide our subjects into three main categories: free riders (F), cooperators (C), and reciprocators (R). However, in order to account for ambiguous or borderline cases we also formed a residual group of 'noisy' players (N) (or, more precisely, subjects with 'noisy' data that are not easily interpretable).

Given that the ultimate goal of our research is to predict behaviour in a repeated PG game, we decided to put more weight on the data collected in the repeated PG game task of Session 1. The classification of players within the repeated PG experiment is a complex matter, as it should take into account various aspects of behaviour. The obvious parameter to start from is the simple average individual contribution over the 20 rounds played 'for real' (IA). This measure of course is the variable under examination and hence is unable to explain itself, but nevertheless can be legitimately used for taxonomic/predictive purposes.[11] A more serious problem is that IA is likely to depend on group behaviour: a subject with an average low contribution, for instance, may be either a pure free rider, or a reciprocator which happens to be in a group of free riders. Similarly, a high average contribution may signal either a pure cooperator, or a reciprocator in a group of reciprocating/cooperating players. Hence, IA needs to be complemented with a second parameter: its difference with respect to average group contribution (DA).

In interpreting PG data we follow a two-steps algorithm. Our algorithm for IA is:

(1) C or R if IA > 12;
(2) R if 12 ≥ IA ≥ 8;
(3) F or R if IA < 8.

Whenever we face an ambiguous case (1 or 3 above), we use the second measure, the difference between individual and group average contribution (AD). The algorithm in this case is:

(4) C if DA ≥ 1;
(5) R if –1 < DA < 1;
(6) F if DA ≤ –1.

The Decomposed Game provides a tight classification of subjects into 5 categories, according to where the 'motivational vector' is placed on the 'value orientation circle'. They are: 'aggressive' if it lies between degree –112.5 and –67.5; 'competitive' if between –67.5 and –22.5; 'individualistic' if between –22.5 and 22.5; 'reciprocating' if between 22.5 and 67.5; 'cooperative' if between 67.5 and 112.5.[12] For our purposes, we aggregated subjects belonging to the first three categories ('aggressive', 'competitive' and 'individualistic') into one single category, that of 'free riders'. Subjects with a low level of coherence (below 50%) were classified as 'noisy'.

[11] In this respect, our research differs from previous attempts to observe individual attitudes or motives in PG-like environments (e.g. Offerman et al. 1996; Brantds and Schram 2001; Andreoni 1995a). Our categories are more 'behavioural' and less psychological in character than those used in those studies.
[12] Notice that here, and everywhere else throughout this paper, our use of the terms 'reciprocator' and 'cooperator' differs from the conventional terminology in the standard literature on the decomposed game. We use the term 'reciprocation' (or 'conditional cooperation') where the standard literature uses the term

The Strategy Method provides two main parameters: the conditional contribution table and the unconditional contribution. We decided to classify as ‘reciprocators’ all players whose conditional contribution functions approximate the ‘perfect reciprocation function’, with a margin of variation of ±10%. For example: a player willing to contribute 0 if the others contribute 0, 0 if they contribute 1, 1 if they contribute 2, 2 if the contribute 3, and so on, would be classified as a reciprocator (albeit with a slight ‘individualistic’ bias). In contrast, a subject willing to give 0 if the others give 0, 0 if they give 1, 0 if they give 2, 0 if they give 3, 1 if they give 4, 2 if they give 5, and so on, would count as a free rider according to our classification. (A symmetrical criterion is used in order to identify cooperators.)

Although the great majority of conditional contribution functions display a coherent pattern (they are either flat or increasing with the group contribution-level), some seem to follow random walks and some other non-easily interpretable patterns.[13] We classified cases like this as ‘noisy’. The conditional contribution table was then compared with the unconditional contribution of the same subject. The idea is that subjects classified as ‘cooperative’ according to the conditional contribution table should give high unconditional contributions, and the opposite should hold for free riders. The behaviour of reciprocators, as usual, is more difficult to pin down, for their unconditional contribution will depend decisively on their expectations about the members of their group. Here we relied on their answer to the question: ‘how much do you think the other members of the group have contributed, on average?’ which was inserted between the unconditional and the conditional contribution tasks. For genuine reciprocators, the answer to the latter question should be highly correlated with the unconditional contribution. Free riders and ‘pure’ cooperators, in contrast, should act pretty much independently of their expectations on others’ behaviour. When these simple rules were violated, the subject was labeled as ‘noisy’.

Finally, we used evidence from the questionnaire to complement the data from the three previous tasks. Questionnaires are not highly valued in experimental economics, but we

‘cooperation’ and we use ‘cooperation’ for behaviour that is there termed ‘altruistic’. This is more in line with the terminology used in the other games.

[13] A significant portion of subjects display puzzling ‘hump-shaped’ functions: their contribution grows with the group contribution up to a point (typically, 50% of the tokens), and then declines towards zero. See also Fischbacher et al. (2001).

found them of some use as an aid in the interpretation of the PG game evidence.[14] The questionnaire included subjects' own *post-hoc* explanation (rationalisation) of the strategies they followed in the repeated PG task. The answers being 'open', the quality of the data varied from subject to subject.[15]

Once this preliminary round of classification had been completed, every single subject was provided with four labels, one for each task. Overall, there was a remarkable degree of convergence between the four classification tasks, but obviously several cases of disagreement (on one or more dimensions) remained. In order to resolve them, we assigned weights to the four classification methods. As anticipated, we gave priority to the PG game data, according to the following formula:

Repeated PG game: 40%
Strategy Method: 20%
Decomposed Game: 20%
Questionnaire 20%.

When no classification reached the 50% level (and therefore in all cases of tie) we assigned the subject to the 'noisy' group (N).

2.6. Example and methodological discussion

In order to give an idea of the classification process as a whole, we shall here illustrate one particular instance taken from our experimental data. The behaviour of subject number 51 in the PG game, vis-à-vis the average contribution level of the members of her group, is represented in Figure 1. Her individual contribution is almost invariably below the group average – Subject 51 was pretty clearly free riding. (Quantitatively, IA = 3.15 and DA = –3.55.) In the Decomposed Game, the subject emerged as a 'competitive' individual (hence a 'free rider' in our terminology), but with a low level of coherence (35%). As a consequence we classified her as 'noisy' in this task. The data from the Strategy Method, conditional contribution task, are represented in Figure 2. The conditional contribution function lies inequivocally below the 45% (or 'perfect

[14] We don't believe, in other words, that questionnaire data can be used as the only or even primary source of evidence, but often they allowed to discriminate between, say, a 'mild' cooperator and a 'noisy' player, or between a cooperator and a reciprocator, etc.

reciprocation') line. This is confirmed by the other two parameters of the Strategy Method: the subject expected other players to contribute on average 100 tokens, but she herself contributed (in the 'unconditional' task) only 50. According to this classification criterion, she is a free rider. Finally, Subject 51 gave the following answer to the question 'What were you trying to achieve in this game?':

> "My main objective was to invest little money in the project so as to obtain a sure gain from the money I had put in the individual account and also to benefit from the gains obtained from the project thanks to the investments made by the other components of the group."

[FIGURES 1 AND 2 ABOUT HERE]

So, to sum up, Subject 51 was classified in the following way:

PG Game: Free rider (40%)
Decomposed Game: Noisy (20%)
Strategy Game: Free rider (20%)
Questionnaire: Free rider (20%)
Aggregate classification: Free rider.

Of course other cases were less straightforward than this one but, as we shall see very shortly, our algorithms seem with hindsight to have performed quite well. This is indeed an important methodological point, which is worth articulating in more detail. As we said in the introduction to this paper, our main objective is to test the 'heterogeneous agents hypothesis' even in the absence of a formal model of reciprocating players. Obviously, working without a formal model means that the classification criteria that we use must necessarily be 'ad hoc' and lack a rigorous theoretical foundation. However, the criteria themselves can receive support from the experimental evidence: if the criteria were totally arbitrary or ill-defined, we should not expect to find any significant difference between the levels of cooperation in Session 1 (with 'heterogeneous' groups) and in Session 2 (with 'homogeneous' groups). To see why, imagine an extreme case in which our classification criteria failed completely: every *true* cooperator, say, would have an equal

[15] Many answers were simply too short or too obscure to be of any use. In order to control as much as possible for our own interpretative bias, we separately evaluated the answers and classified subjects

chance of being assigned to any of the four groups (F, C, R, N). This would amount to a mere random reallocation of the subjects to different groups, and therefore we should expect in Session 2 to observe the same phenomenon of decaying overcontribution as in Session 1. In contrast, the stronger the difference between the two sessions, the more our (admittedly ad hoc) classification criteria turn out to be corroborated.

Another piece of evidence corroborating the classification criteria is the convergence of their results. As microbiologists know very well, it is not necessary that we are 100% confident in the reliability of a microscope in order to obtain valid empirical data. We can use different instruments (e.g., a light microscope, an electronic microscope, etc.) to observe the same specimen and then use an argument from coincidence of the following sort: it would be a true miracle if all these different instruments reported the same, mistaken, phenomenon. Their convergence suggests that the phenomenon must be real, rather than an artefact of the observation procedure. Our use of different classification techniques followed the same logic (a procedure sometimes called ‘triangulation’).

As a matter of fact, we did observe a highly significant difference in the contribution patterns of Sessions 1 and 2 in our experiment (as we shall see in detail in the next session); and we did observe a remarkable convergence between the different methods of classification. Hence, we can conclude with a high degree of confidence that the classification procedure did the job.

2.7. Session 2: Repeated linear PG experiment (homogeneous groups)

All subjects were classified during the week that elapsed between Session 1 and Session 2. The number of subjects falling in each category is shown in Table 1.

	Free riders	Reciprocators	Cooperators	Noisy	Total
N	29	32	17	14	92

Table 1: Classification

As expected the numbers in each category are not multiples of four, which caused some inconvenience in creating new homogeneous groups. In order to achieve maximum

accordingly; then, we compared our results and resolved the (rare) cases of disagreement.

homogeneity and not to lose too many subjects, we devised a system of substitutes to replace subjects who might unexpectedly fail to show up at Session 2. Eventually, we managed to create 6 groups of free riders, 6 groups of reciprocators, 4 groups of cooperators, and 3 groups of 'noisy' players. We also formed two non-homogeneous groups by matching the remaining players. One group turned out to be made of three reciprocators and one free rider (as we shall see, the results from this group are quite surprising); the other group of one free rider, two reciprocators, and one 'noisy' player. For obvious reasons, these non-homogenous groups will not be considered in the main process of data-analysis below. Eight subjects (three free riders, three reciprocators, one cooperator, and one 'noisy' player) did not participate in Session 2, either because they failed to show up, or because there were no other players available to match them with.[16]

Three days after the first session subjects were contacted individually (by email) and confirmed the exact time of the second experimental session. When they arrived, subjects were allocated individually to their own terminal. The instructions were identical to those of session 1, task 3, except that it was made clear that the composition of the groups differed from that of the first session. Once again, we let them play 3 rounds for training and 20 for real. At the end of the experiment we added their earnings from both sessions, converted into Euros, and paid them accordingly.

3. Results

The core of our analysis is based on the behaviour of 64 subjects: 4 groups of cooperators, 6 of free riders, and 6 of reciprocators. (Towards the end of this section, however, we shall also comment briefly on the behaviour of 'noisy' players and on the two non-homogeneous groups.) The main predictions of the heterogeneous agents hypothesis turn out to be corroborated: (1) in the second PG game the average level of contribution is significantly higher than in the first one; (2) in the second PG game groups of ('pure' and 'conditional') cooperators display a very high and fairly constant level of contribution. In groups of free riders, in contrast, the contribution starts lower and quickly jumps to a low level, reaching in the last round the expected Nash equilibrium of zero contribution.

[16] The subjects who did not show up in both sessions were not paid at all. Those who did show up but could not participate in the experiment in Session 2, received their earnings from Session 1, plus a flat fee equal to

[FIGURE 3 ABOUT HERE]

The first result is apparent from Figure 3: the average contribution level in Session 2 is always above that of Session 1. Statistically, the hypothesis that the two sets of data (contribution in the PG Game, Session 1, and contribution in the PG Game, Session 2) come from the same population is rejected at the 1% level in all rounds but six, where it is rejected at 5% (Wilcoxon signed ranks test). The question then arises of which subjects effectively contributed to the upward shift of the contribution curve. Table 2 show that reciprocators are mostly responsible for the shift: on average, their contribution moved from 10.4 to 18.6 tokens. This abrupt increase is to be expected in the light of the heterogeneous agents hypothesis. What the hypothesis does not necessarily imply is that cooperators should also raise their contribution level once placed in homogenoeus groups (+1.5), nor that free riders should end up free riding more (–2.3). One plausible explanation for this is that our groups of cooperators and free riders were 'infiltrated' by a certain amount of reciprocators – i.e. they were less 'pure' than we wanted them to be. In the case of free riders, it is also likely that in the first session some of them played strategically, trying to profit by contributing less than the group average but still above the Nash equilibrium (in order not to discourage others' contributions too early in the game). In the second session, they might have quickly realised that such a strategy was not profitable, and provoked a more rapid decay.

	Session 1	Session 2
Cooperators (N=16)	14.91	16.45
Reciprocators (N=24)	10.38	18.65
Free riders (N=24)	5.03	2.71

Table 2: Average contribution in the repeated PG game

Figures 4 (a, b, c), 5 and 6 show the average contribution of each type of player during the 20 rounds of play, in Sessions 1 and 2. Notice how in Session 1 (Figure 5) the contribution level of reciprocators is placed between that of free riders and cooperators, but then shifts up to the top of the graph in Session 2 (Figure 6). Part of the explanation is that in Session 2 the variance is greater among cooperators (which is probably due to the

---

the average earnings of the other players in Session 2.

fact that we were not terribly successful in forming homogeneous groups for this category).

[FIGURES 3 (a, b, c), 4 AND 5 ABOUT HERE]

As was to be expected, in Session 2 free riders start with a lower level of contribution, which tends to decay very quickly (the average is 5 tokens after only two rounds) and never manages to recover. Cooperators and reciprocators start with a very high level of contribution, which remains high throughout most of the game. The behaviour of reciprocators is particularly impressive, with constant (and almost full) contribution until round 19. There is evidence of an 'end-effect' (Andreoni, 1988; Keser, 1996) in the last 3-4 rounds, with abrupt decay to a level of about 12 tokens. (Such an 'end-effect', incidentally, explains why the difference between the contribution levels in the first and second session is greater in the middle of the game – cf. Figure 3.) The presence of an end effect in reciprocating players casts further doubts on the explanations based on pure altruism. Many reciprocating players do defect at the end of the game, which suggests that their behaviour is indeed prompted either by selfish motives or at least by fear of being exploited. It is likely that reciprocators do not consider the equilibrium strategy to be rational (they probably cannot bring the backward induction argument to its radical and counterintuitive consequences.)

We ran a Kruskal-Wallis non-parametric test on the contribution in the second session of the four samples (free riders, reciprocators, cooperators, and 'noisy'players). The results, summarised in Table 3, show that the hypothesis that the four samples of data come from the same generation mechanism can be rejected with near certainty. We also ran Mann-Whitney tests on pairs (free riders vs. cooperators, cooperators vs. reciprocators, etc.), and the results are as expected (see Table 4). The difference between free riders and, respectively, cooperators and reciprocators, is highly significant (it passes the test at the 1% level in *all* rounds). As expected, the difference in behaviour of cooperators and reciprocators in the second session is small and the two groups cannot be easily distinguished .

Rd1	rd2	rd3	rd4	rd5	rd6	rd7	rd8	rd9	rd10	rd11	rd12	rd13	rd14	rd15	rd16	rd17	rd18	rd19	rd20
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Table 3: Kruskal Wallis test on all groups (F, C, R, N), Session 2; asymptotic significance.

	rd1	rd2	rd3	rd4	rd5	rd6	rd7	rd8	rd9	rd10	rd11	Rd12	rd13	rd14	Rd15	rd16	rd17	rd18	rd19	rd20
C vs F	.000	.000	.000	.000	.000	.000	.000	.003	.000	.001	.000	.000	.000	.000	.000	.000	.000	.000	.002	.002
C vs R	.467	.570	.968	.885	.682	.517	.644	.375	.435	.435	.500	.781	.404	.781	.552	.926	.227	.295	.188	.721
F vs R	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Table 4: Mann Whitney test, Session 2; exact significance.

We haven't said much about 'noisy' players. In fact, it is difficult to characterise them except for the fact that their choices show a much greater variance than those of the other groups[17]. On average their contribution levels are above those of free riders in both sessions, and below those of cooperators. This is not surprising, given that the 'noisy' group was used as a 'catch-all' and therefore probably included players of all types. The two non-homogeneous groups engaged in fairly strange behaviour, which we cite only for curiosity and the sake of completeness. The group composed of three reciprocators and one free rider managed to sustain a decent level of cooperation – below the average of reciprocators but well above that of free riders. The group composed of two reciprocators, one noisy player and one cooperator displayed an even lower, and swinging, level of cooperation throughout the game.

4. Conclusion

The altruism plus error hypothesis implies that repeated play should decrease contribution, but this does not seem to be true. Or rather: it is true for free riders and for reciprocators in non-homogeneous groups, but certainly not for cooperators and reciprocators in homogeneous groups. Our evidence strongly supports the heterogeneous agents hypothesis. This may not be relevant in *all* economic contexts, but it certainly is in some (Heckman 2001a, 2001b). In particular, representative agent models (especially the 'augmented' ones with altruism[18]) may be perfectly adequate in many circumstances for

[17] In the last few rounds the variance increases also among cooperators and reciprocators (the end-effect is probably involved here), while it remains small and decreases to zero in the last round among free riders.
[18] See Fehr and Schmidt (1999), Bolton and Ockenfels (2000).

predictive purposes, but lack explanatory depth and might fail to capture some important mechanisms that sustain cooperation. This should have important consequences on the experimental debate on repeated PG games. The divide between models of self-interested agents on the one hand and models of altruistic players on the other, might never be resolved simply because there are agent of both types. It seems more interesting and fruitful to recognise not only the existence of these two types of players, but also their influence on another (large) category of players: reciprocators (or 'conditional cooperators').

Since reciprocators constitute a large portion of the experimental population (at least in our sample, but see also Fischbacher et al., 2001)[19], it is possible to raise the overall level of contribution by forming homogeneous groups of players with similar attitudes towards cooperation. Letting people choose their partners may be a key to the promotion of cooperating behaviour.

References

Anderson, S.P., J.K. Goeree and C.A. Holt (1998) "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games", *Journal of Public Economics*, 70, 297-323.

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

Andreoni, J. (1995a) "Warm Glow vs. Cold Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments", *Quarterly Journal of Economics*, 110, pp. 1-21.

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

Brandts, J. and A. Schram (2001) "Cooperation and Noise in Public Goods Experiments: Applying the Contribution Function Approach", *Journal of Public Economics*, 79, 399-427.

Burlando, R.M. and P. Webley (1999) "Individual Differences and Long-run Equilibria in a Public Good Experiment", in *Inquiries into the Nature and Causes of Behavior*: Proceedings of the 24th IAREP Annual Colloquium, Belgirate, Italy.

[19] In this experiment we probably underestimated the number of reciprocators, but it is always difficult to find 'pure' types, especially at the extremes of the distribution. Certainly it should be possible to distinguish among reciprocators with a more 'altruistic' attitude and others that are simply fearful of being exploited or are even inclined to free riding. We attempted a more detailed analysis, identifying three sub-types for each category of players, but we had not enough participants to allow us to do this properly. And the pay off of finer distinctions is quite dubious.

Fehr, E. and U. Fishbacher (2002) “Why Social Preferences Matter – The Impact of Non-selfish Motives on Competition, Cooperation and Incentives”, *Economic Journal*, 112, C1-C33.
Fehr E. and K.M. Schmidt (1999) “A Theory of Fairness, Competition, and Cooperation”, *Quarterly Journal of Economics*, 114, 817-868.
Fischbacher, U., S. Gächter, and E. Fehr (2001) “Are People Conditionally Cooperative? Evidence from a Public Goods Experiment”, *Economics Lette*rs, 71, 397-404.
Griesinger, D.W. and J.W. Livingston (1973) “Towards a Model of Interpersonal Motivation in Experimental Games”, *Behavioral Science*, 18, 173-188.
Heckman, J.J. (2001a) “Accounting for Heterogeneity, Diversity and General Equilibrium in Evaluating Social Programmes”, *Economic Journal*, 111, F654-699.
Heckman, J.J. (2001b) “Microdata, Heterogeneity, and the Evaluation of Public Policy”, *Journal of Political Economy.*
Keser, C. (1996) “Voluntary Contributions to a Public Good When Partial Contribution is a Dominant Strategy”, *Economics Letters*, 50, 359-366.
Ledyard, J.O. (1995) “Public Goods: A Survey of Experimental Research”, in J. Kagel and A. Roth (eds.) *The Handbook of Experimental Economics*, Princeton: Princeton University Press, pp. 111-194.
Levine, D. (1998) “Modelling Altruism and Spitefulness in Experiments”, *Review of Economic Dynamics*, 1, 593-622.
Liebrand, W.B.G. (1984) “The Effect of Social Motives, Communication and Group Size on Behavior in an N-person Multi-stage Mixed-motive Game”, *European Journal of Social Psychology*, 14, 239-264.
Johnson, D.W. and A.A. Norem-Hebeisen (1979) “A Measure of Cooperative, Competitive and Individualistic Attitudes”, *Journal of Social Psychology*, 109, 253-261.
Offerman, T., J. Sonnemans, and A. Schram (1996) “Value Orientations, Expectations, and Voluntary Contributions in Public Goods”, *Economic Journal*, 106, 817-845.
Palfrey, T.R. and J.E. Prisbey (1996) “Altruism, Reputation and Noise in Linear Public Goods Experiments”, *Journal of Public Economics*, 61, 409-427.
Palfrey, T.R. and J.E. Prisbey (1997) “Anomalous Behavior in Public Goods Experiments: How Much and Why?”, *American Economic Review*, 87, 829-846.
Plott, C.R. (1995) “Rational Individual Behaviour in Markets and Social Choice Processes: The Discovered Preference Hypothesis”, in K.J. Arrow, E. Colombatto, M. Perlman and C. Schmidt (eds.) *The Rational Foundations of Economic Behaviour*, London: Macmillan, pp. 225-250.
Rabin, M. (1993) “Incorporating Fairness into Game Theory and Economics”, *American Economic Review*, 83, 1281-302.
Sugden, R. (1984) “Reciprocity: The Supply of Public Goods Through Voluntary Contributions”, *Economic Journal*, 94, 772-787.
Webley, P., D.J. Hessing, H.S.J. Robben and H. Elffers (1988). “Social Orientations in Social Dilemmas: Explorations in Operationalisations”, Paper presented at the 3rd Conference on Social Dilemmas, Groningen.
Weiman, J. (1994) “Individual Behaviour in a Free Riding Experiment”, *Journal of Public Economics,* 54, 185-200.

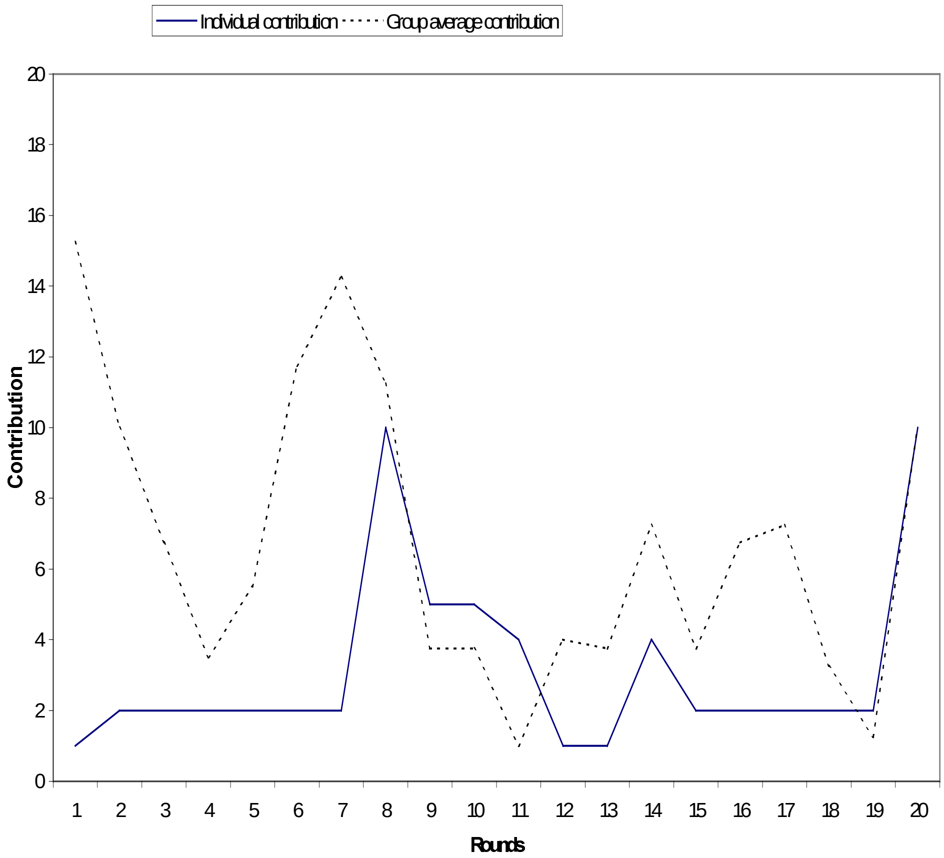


Figure 1: Subject 51, Session 1, PG game

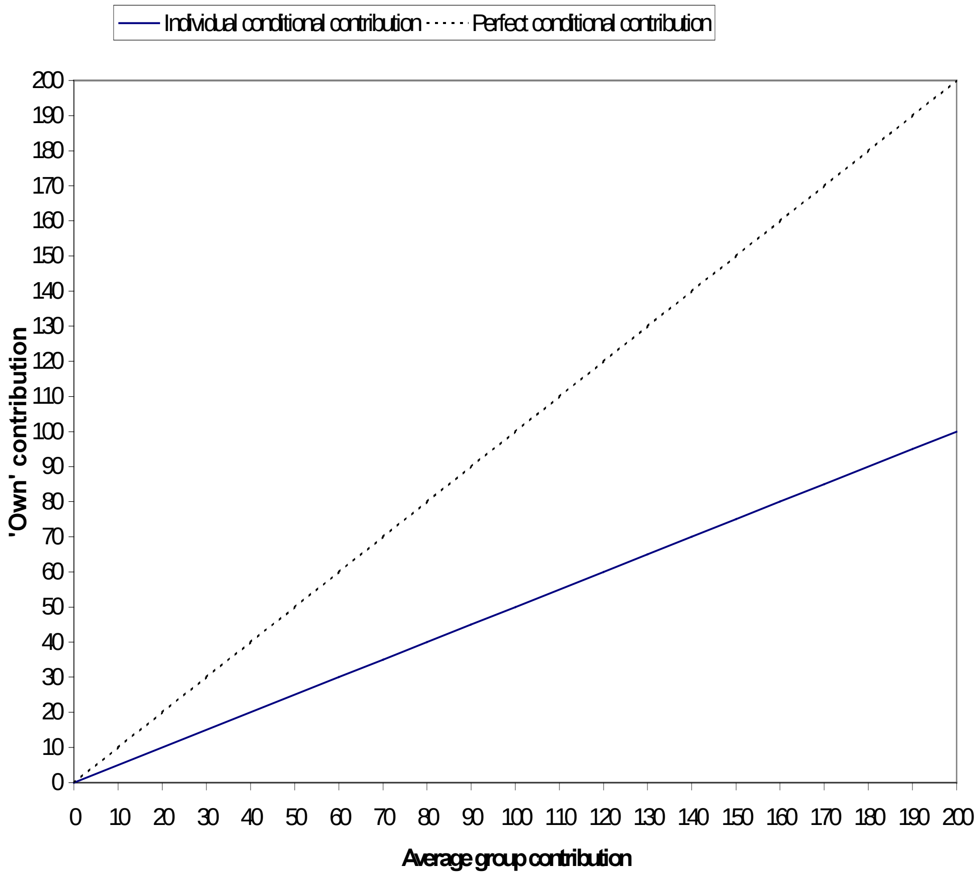


Figure 2: Subjet 51, Session 1, Strategy Method

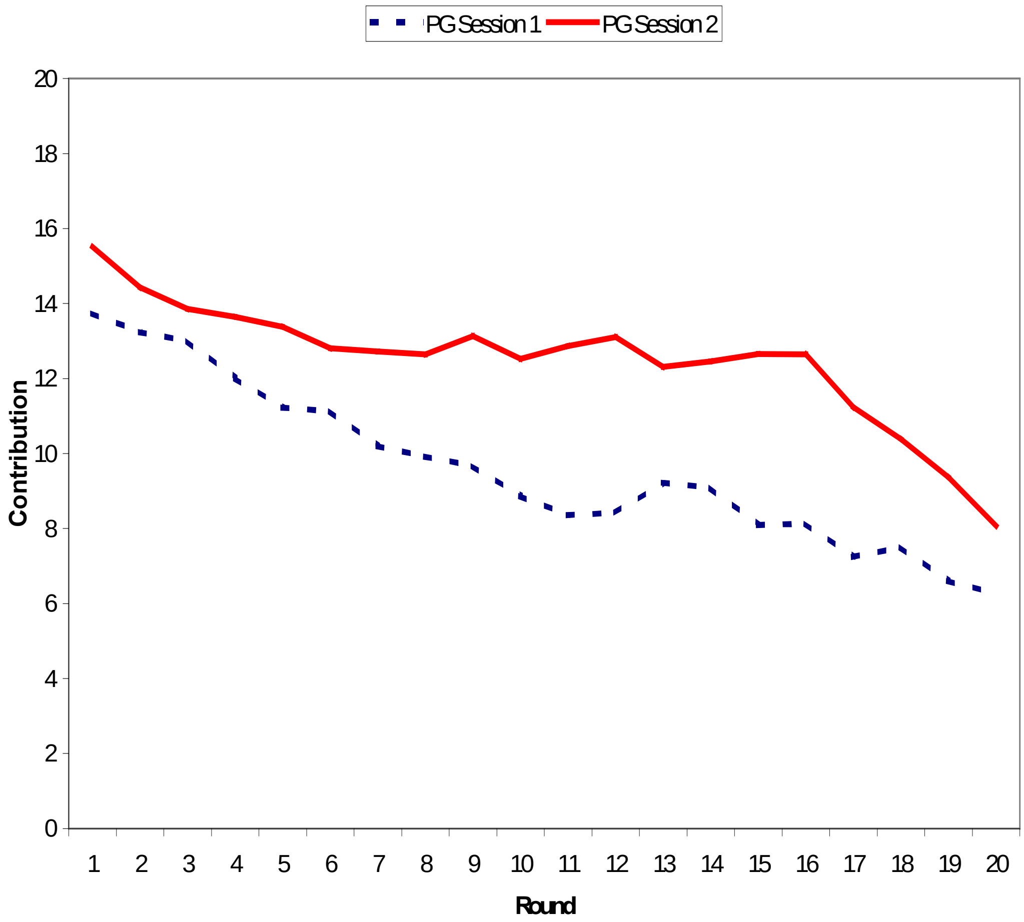


Figure 3: Contribution level in the first and second PG game

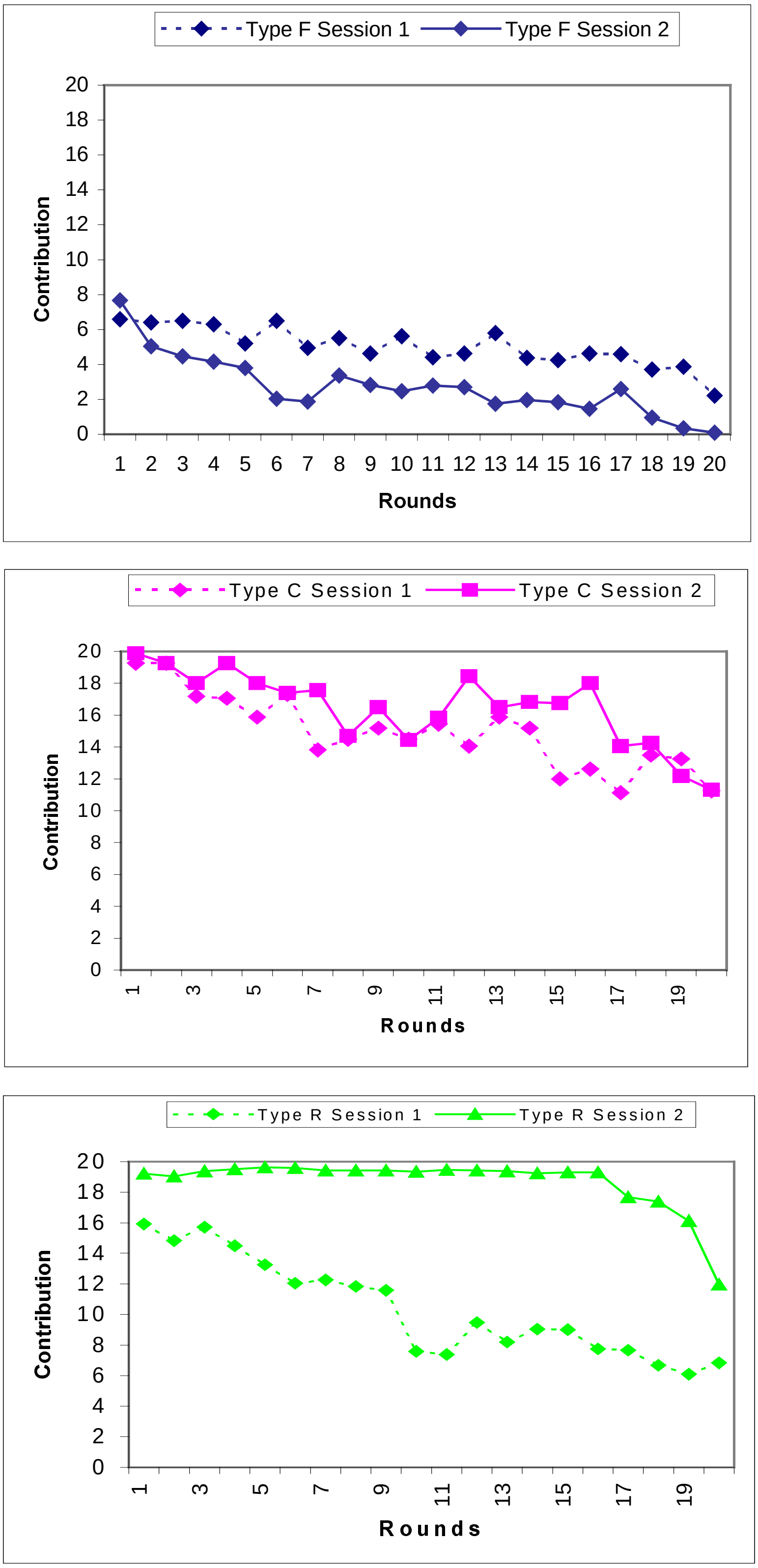


Figure 4 (a,b,c): Contribution in PG Game, Sessions 1 vs. 2, per type

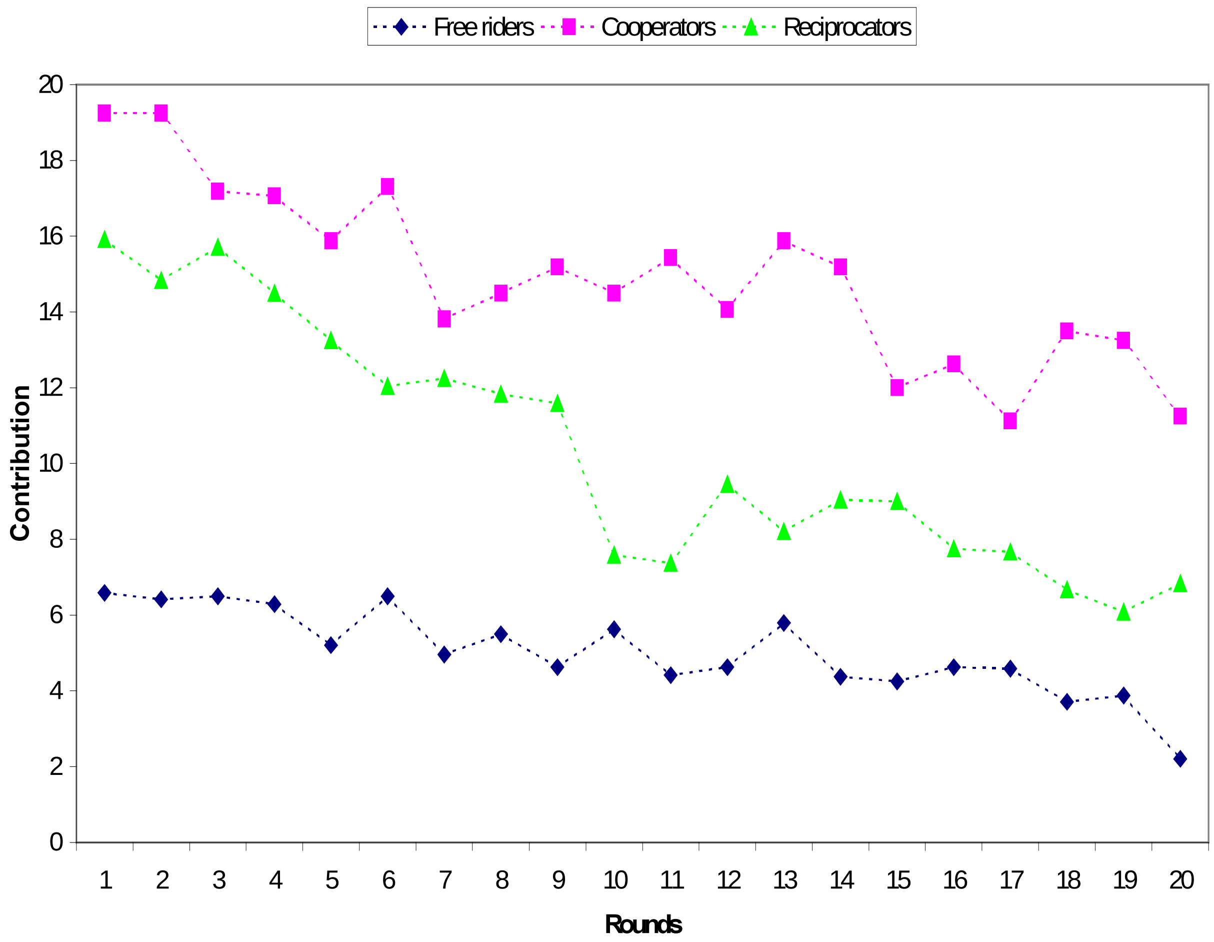


Figure 5: Contribution level by type of player: first session

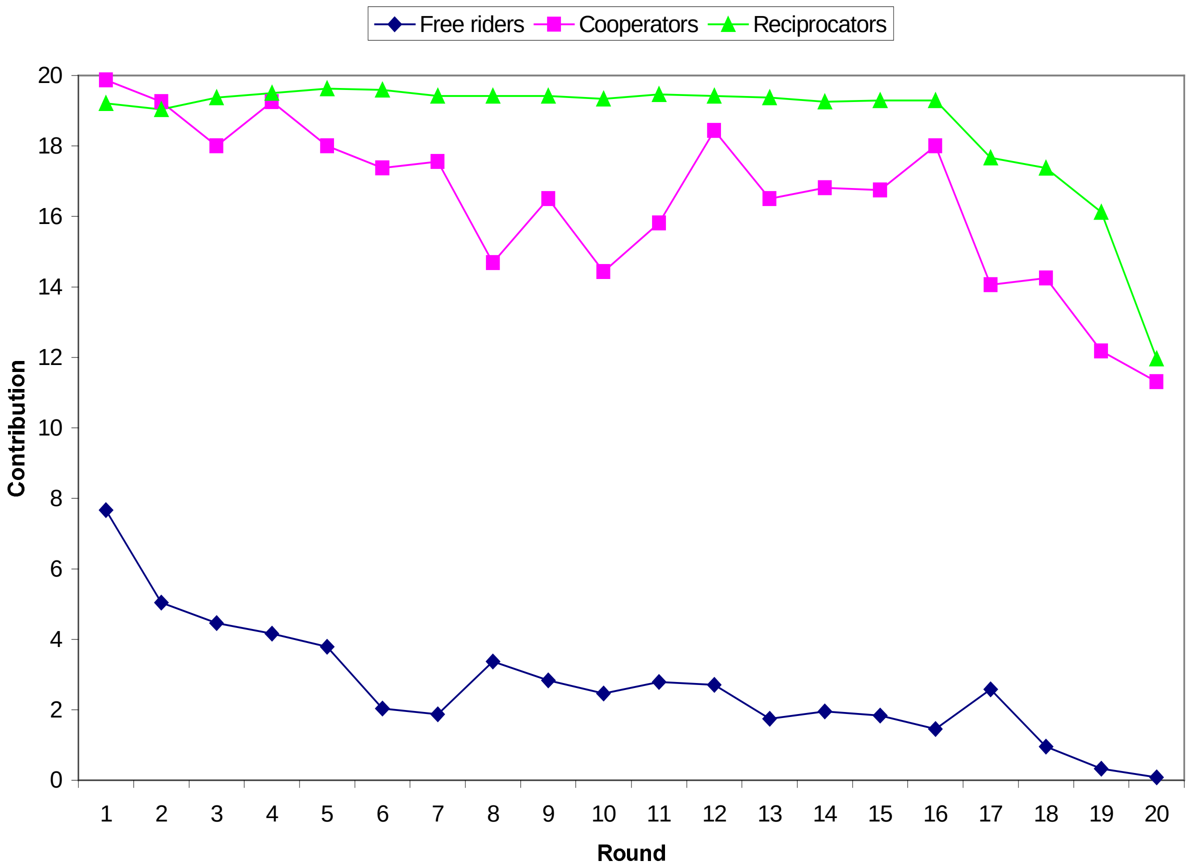
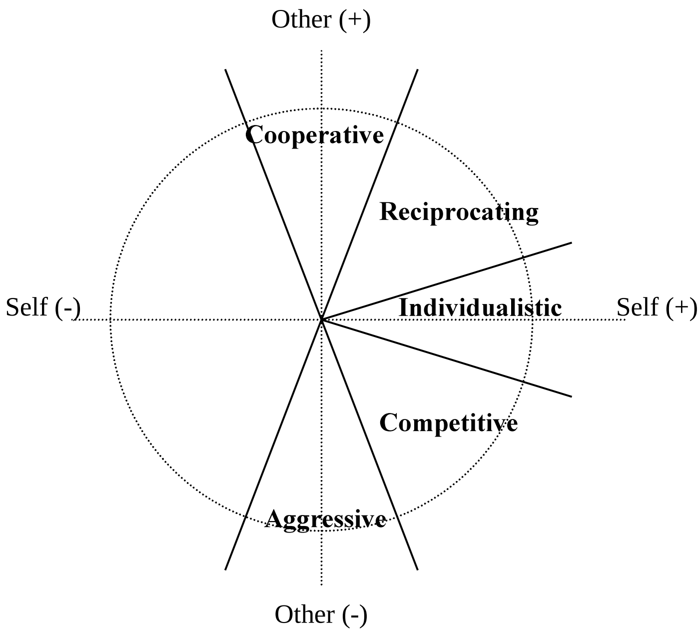


Figure 6: Contribution level by type of player: second session

Appendix 1

	Option A		*Option B*	
Question	***Self***	***Other***	***Self***	***Other***
1	+150	0	+145	+39
2	+144	-39	+130	-75
3	+130	-75	+106	-106
4	+106	-106	+75	-130
5	+75	-130	+39	-145
6	+39	-145	0	-150
7	0	-150	-39	-145
8	-39	-145	-75	-130
9	-75	-130	-106	-106
10	-106	-106	-130	-75
11	-130	-75	-145	-39
12	-145	-39	-150	0
13	-150	0	-145	+39
14	-145	+39	-130	+75
15	-130	+75	-106	+106
16	-106	+106	-75	+130
17	-75	+130	-39	+145
18	-39	+145	0	+150
19	0	+150	+39	+145
20	+39	+145	+75	+130
21	+75	+130	+106	+106
22	+106	+106	+130	+75
23	+130	+75	+145	+39
24	+145	+39	+150	0

Decomposed Game: choice of allocations



Decomposed Game: the 'value orientation circle'