Human Cooperative Behavior

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität

zu Kiel

vorgelegt von

Dirk Semmann

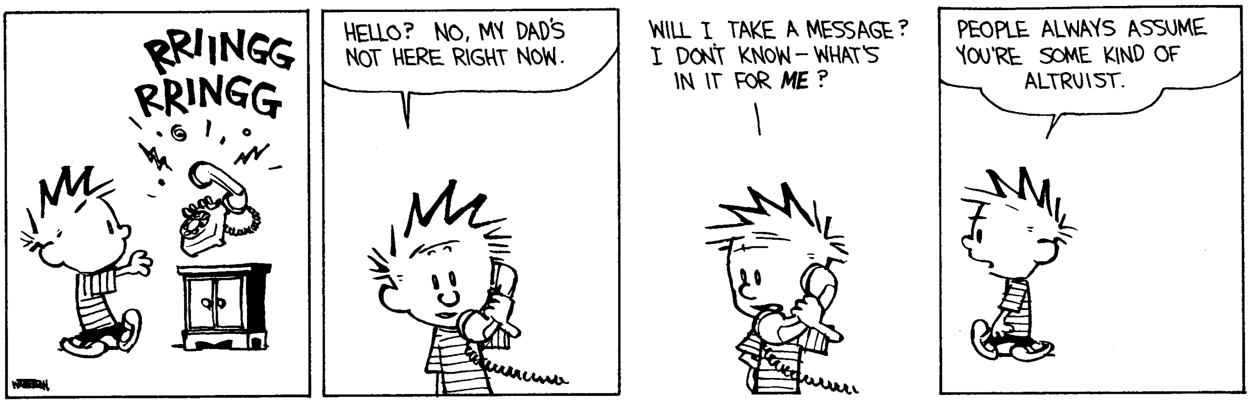
Kiel 2003

Referent: Prof. Dr. M. Milinski

Korreferent: Prof. Dr. Heinz Brendelberger

Tag der mündlichen Prüfung: 16.02.2004

Zum Druck genehmigt: 16.02.2004



Bill Watterson (1996)

Contents

Summary

Evolutionary theory provides the biological sciences, with a fundamental and powerful model to explain the emergence of cooperative behavior. A detailed explanation for the existence of cooperation between related individuals is provided by the theory of kin selection. When kin cooperate the helper gives the receiver an advantage and thereby increases the relative probability that copies of his own genes are present in the next generations. However, one cannot explain examples of apparent altruism through kin selection, because in these cases unrelated individuals interact. The answer for many of these examples is provided by the theory of reciprocal altruism, where individuals behave reciprocal by returning help to a previous donor. By applying these two theories it is possible to explain many, but by far not all cooperative situations. There have to be other mechanisms that lead to cooperation and sustain already established cooperation.

In my dissertation I have tested empirically new models and predictions of how cooperation between unrelated humans can be established. This research is especially important because we interact in a close net of relationships, where cooperation between unrelated individuals plays one of the main roles. Modern human societies are impossible to imagine without cooperation between unrelated individuals. By identifying the circumstances under which cooperation is stable between unrelated individuals, it will be possible to understand the deciding factors in politics, economy and in our private lives. As a consequence we would be provided with intellectual tools to positively influence the deciding factors by alternating the circumstances accordingly.

We are often not aware of the importance of cooperation between unrelated partners in our daily lives. Regularly people find unconsciously cooperative solutions, for instance when they try simultaneously to walk through a narrow door. Some professions depend very strongly on cooperative behavior between unrelated colleagues. To act uncooperatively in such a profession can endanger the health or even the lives of the colleagues (e.g. firemen and firewoman). Cooperative strategies for these kinds of situations have to have evolved and need to be evolutionary stable, otherwise we would hardly ever find cooperative behavior in the present and

then only between related individuals. According to the evolutionary theory the cooperative strategies found today, also have to provide an advantage to the bearer.

For a long time economists and biologists have been interested in the emergence and sustainability of cooperative behavior. Nevertheless, only with the introduction of game theory, a mathematical basis was established to incorporate this behavior into biological evolutionary models. From then on it was possible to make predictions with the help of theoretical models, about the circumstances under which certain behavioral patterns emerge and what underlying mechanisms possibly sustain these patterns.

In my dissertation I have empirically tested predictions of circumstances that promote cooperative behavior between unrelated humans. The main results of my work are the following: (i) Humans often donate money to charity. On first sight this seems to be a disadvantageous behavioral trait. Donations to charity include costs that reduce the direct fitness of the individual. However, it has to be beneficial to the bearer, otherwise it would be eliminated from the population through evolutionary processes. The study showed that there is indeed an advantage. By donating money to charity (here to UNICEF) one builds up a good personal reputation in the own social group. Participants that donated to UNICEF received with higher probability help from other participants and were as well more likely to be voted the group representative. (ii) Reputation is known to be an important currency in indirect reciprocity games. Humans therefore should also try to establish a good reputation in other social games, when this reputation is known in future indirect reciprocity games. Humans are in general unable to sustain a public resource that everybody is free to overuse anonymously. Is it possible that humans sustain a public resource if the use of the resource is linked to the personal reputation? The experiment showed, that the risk of loosing a good reputation by overusing the public resource actually lead to sustaining it. Furthermore the public resource was not only sustained, but also turned out to be surprisingly profitable to all group members. (iii) A theoretical model supplied a new possibility to sustain a public resource and hereby make humans act cooperatively. The strikingly simple idea was to introduce the possibility not to participate in the public goods group and instead use a personal resource with a low but sure payoff. The prediction was an always recurring rise to dominance, of three strategies ((a) not to participate in the public goods group, (b) participate in the public goods group and to cooperate within the group or (c) participate and to defect within the group) within the population. This dynamic was expected because whenever most members of the population choose the same strategy, one of the other two strategies had a higher payoff. The same type as the predicted dynamic has also been found in models of the famous children game of “rock-paper-scissors”. The model predicted that the public resource is sustained by the ongoing dynamic, which is liked with a recurring rise of cooperative behavior. Is it enough to supply humans with the possibility not to participate in the public goods game to produce such recurring rise of cooperation? The dynamic was established as predicted, whereby the changing dominance of the three strategies with repeated cooperative phases could be observed and the resource was on average sustained. (iv) When humans make decisions about using a public resource, which at some times are reputation relevant and at other times are not reputation relevant, do they use this information strategically? In this study it was shown that, humans are aware when their decisions are not reputation relevant and immediately reduce their cooperation to maximize their personal profit. Once more, as soon as the decisions about using the public resource were linked to the reputation, cooperation was much higher and the resource was sustained. (v) In some potentially cooperative situations humans’ meet partners from outside the own social group. These “strangers” have a reputation that they have built in another social group. Do humans put a different value on a strangers’ reputation in comparison to the reputation of members of the own social group? It was shown that it is not relevant if the reputation was built within the own or in a foreign social group.

In summary we found the following: Humans behave uncooperatively, when it is to the personal advantage. However, certain circumstances lead to cooperative behavior in humans. Reputation building is one of the most important mechanisms in this context, which enables us to cooperate even with not related strangers. However humans consciously make strategic use of situations where they do not harm their reputation by behaving uncooperatively. Nevertheless, even in completely anonymous situations it is possible to create circumstances, like introducing optional participation in the public goods situations, which promote cooperative behavior in humans.

Zusammenfassung

Die Biologie liefert mit der Evolutionstheorie ein starkes grundlegendes Erklärungsmodell für das Entstehen von kooperativem Verhalten. Kooperatives Verhalten zwischen verwandten Individuen kann detailliert durch die Theorie der Verwandtenselektion erklärt werden. Helfen Verwandte einander, verschafft der Helfer dem Empfänger einen Vorteil und erhöht damit die relative Wahrscheinlichkeit, daß Kopien seiner eignen Gene in den nächsten Generationen häufiger vertreten sein werden. Durch Verwandtenselektion konnten jedoch Beispiele von scheinbarem Altruismus nicht erklärt werden, da diese zwischen nicht verwandten Individuen statt finden. Die meisten dieser Beispiele ließen sich erst mit der Formulierung der Theorie des reziproken Altruismus erklären. Wird einer Person geholfen, und die Hilfe wird vom Empfänger später erwidert, dann spricht man von reziprokem Verhalten. Doch welche anderen Mechanismen führen zu Kooperation und erhalten Kooperation in einem bestehenden System?

In meiner Dissertation habe ich neue Modelle zu kooperativem Verhalten zwischen nicht verwandten Individuen beim Menschen empirisch getestet. Dies ist besonders wichtig, weil wir uns in einem Beziehungsnetz bewegen, in dem Kooperation zwischen nicht verwandten Partnern eine zentrale Rolle spielt. Moderne, menschliche Gesellschaften sind ohne die Kooperation zwischen nicht verwandten Individuen nicht mehr vorstellbar. Wenn die Rahmenbedingungen für Kooperation zwischen nicht verwandten Menschen gefunden worden sind, wird man in der Lage sein Entscheidungsmuster in der Politik, Wirtschaft und im privaten Leben zu verstehen und positiv zu beeinflussen. Die Beeinflussung erfolgt indem man die Rahmenbedingungen anpaßt.

Die zentrale Rolle der Kooperation zwischen nicht verwandten Partnern im menschlichen Alltag ist uns häufig nicht bewußt. Dieses geschieht z.B. wenn zwei Personen gleichzeitig auf eine schmale Tür zugehen. Hier kommt es zwischen beiden Personen in aller Regel zu einer kooperativen Einigung, wer zuerst durch diese Tür gehen darf. In einigen Berufen, wie z.B. bei der Feuerwehr, kann unkooperatives Verhalten durch die nicht verwandten Kollegen jemanden das Leben oder die Gesundheit kosten. Für diese Situationen zwischen nicht verwandten Individuen muß es Strategien geben, welche evolutionär stabil sind. Andernfalls

würden wir kooperatives Verhalten heutzutage nur sehr selten und fast ausschließlich zwischen verwandten Individuen finden. Dies bedeutet, daß es Strategien geben muß die ihrem Träger, indem er sich kooperativ verhält, einen Vorteil verschaffen.

Ökonomen und Biologen sind seit langem sehr an der Entstehung und der Stabilität von kooperativem Verhalten interessiert. Erst mit der Einführung der Spieltheorie bot sich eine mathematische Grundlage, um diese Form von Kooperation in die bestehenden evolutionsbiologischen Modelle zu integrieren. Mit Hilfe der Spieltheorie konnten nun Vorhersagen gemacht werden, unter welchen Bedingungen bestimmte Verhaltensweisen entstehen können, und durch welche grundlegenden Mechanismen diese Verhaltensweisen erhalten bleiben.

In dieser Dissertation habe ich mit Hilfe spieltheoretischer Modelle, empirisch die Randbedingungen untersucht, welche kooperatives Verhalten zwischen nicht verwandten Menschen fördern. Die Haupterkenntnisse meiner Arbeit sind die Folgenden: (i) Menschen spenden häufig zu wohltätigen Zwecken. Dabei entstehen ihnen Kosten und damit reduzieren sie ihre persönliche Fitneß. Es erscheint auf den ersten Blick zum Nachteil eines Trägers dieses Verhaltensmerkmals. Diese Verhaltensweise muß jedoch zum Vorteil des Spenders sein, da ansonsten evolutionäre Prozesse dafür sorgen würden, daß dies Verhalten aus einer Bevölkerung eliminiert wird. Die Studie zeigte, daß der Vorteil darin besteht, daß die persönliche Reputation innerhalb der eigenen sozialen Gruppe durch Spenden (in unserem Experiment an UNICEF) steigt. Es stellte sich heraus, daß Personen, die viel an UNICEF gespendet haben, mit höherer Wahrscheinlichkeit Hilfe von anderen bekamen und bevorzugt zu Repräsentanten der Gruppe gewählt wurden. (ii) Reputation ist eine wichtige Währung bei indirekter Reziprozität. Daher ist es wahrscheinlich, daß Menschen auch in anderen sozialen Interaktionen Reputationspflege betreiben, wenn die hier geschaffene Reputation auch in indirekten Reziprozitätssituationen bekannt ist. Normalerweise werden limitierte Gemeinschaftsressourcen, die ohne Beschränkungen der Allgemeinheit zugänglich sind, von Menschen übernutzt. Ist es möglich, diese Gemeinschaftsressourcen zu erhalten, in dem man die Nutzung einer solchen Ressource mit der persönlichen Reputation verknüpft? Es zeigte sich, daß die Gefahr des Reputationsverlustes durch eine Übernutzung der Gemeinschaftsressource tatsächlich dazu führte, daß diese erhalten wurde. Die Gemeinschaftsressource wurde nicht nur erhalten, sondern sie war überraschend produktiv für die Gruppenmitglieder. (iii) Ein theoretisches Modell beinhaltete eine neue Möglichkeit eine Gemeinschaftsressource zu erhalten und Menschen dazu zu bewegen, miteinander zu kooperieren. Die Möglichkeit sich nicht an der Gemeinschaftsressource zu beteiligen, sondern statt dessen eine individuelle, nicht so ergiebige Ressource zu nutzen, sollte dazu führen, daß ein immerwährender Dominanzwechsel zwischen drei Strategien ((a) nicht an der Gruppe teilnehmen, die die Gruppenressource nutzen kann, (b) teilzunehmen und nicht zu kooperieren oder (c) teilzunehmen und zu kooperieren) in einer Population entsteht. Dieser Effekt entsteht dadurch, daß sobald sich die Mehrzahl der Individuen für eine bestimmte Strategie entscheidet, der mögliche Profit höher bei einer der beiden anderen Strategien ist. Diese Dynamik entspricht der des bekannten Kinderspiels „Stein-Schere-Papier“. Durch die anhaltende Dynamik und die immerwiederkehrende Kooperation wird die Ressource erhalten. Reicht es aus, Menschen die Möglichkeit zu geben, nicht an der Gruppenressource teilzunehmen, um Kooperation entstehen zu lassen? Die Dynamik entstand wie vorhergesagt. Es gab einen steten Dominanzwechsel der drei Strategien in der Gruppe, womit sich immer wieder kooperative Phasen ergaben und die Ressource erhalten wurde. (iv) Wird in einer Situation, in der zeitweilig die Entscheidung eine Gemeinschaftsressource zu nutzen mit Reputation verknüpft ist, es zu anderen Zeiten aber möglich ist zu entscheiden ohne die Reputation zu schädigen, dieses Wissen strategisch verwendet? In dieser Studie wurde festgestellt, daß Menschen sehr bewußt Reputationsirrelevante Situationen für sich nutzen und sich hier zu ihrem persönlichen Vorteil unkooperativ verhalten. Waren die Entscheidungen jedoch mit Reputation verknüpft, wurden die Entscheidungen sehr viel kooperativer, und in diesen Phasen wurde wie zuvor die Gemeinschaftsressource nicht übernutzt. (v) Es gibt Situationen in denen Menschen auf mögliche Kooperationspartner treffen, die aus einer anderen Gemeinschaft kommen. Diese „Fremden“ haben in ihrer eigenen Gruppe eine Reputation aufgebaut. Wird diese Reputation anders bewertet als die Reputation von Personen aus der eigenen Gemeinschaft? Es konnte gezeigt werden, daß es nicht entscheidend ist, ob ein Partner seine Reputation in der eigenen Gruppe oder in einer fremden Gruppe aufgebaut hatte. Es ist nur nötig, daß die Information über die Reputation zur Verfügung steht.

Zusammenfassend ergibt sich folgendes Bild: Wenn es zu ihrem persönlichen Vorteil ist, verhalten sich Menschen unkooperativ. Es gibt jedoch Randbedingungen die menschliche Kooperation fördern. Reputation ist ein sehr wichtiger Mechanismus in diesem Zusammenhang. Reputationsbildung ermöglicht es uns auch mit fremden, nicht verwandten Personen tagtäglich zu kooperieren. Menschen nutzen strategisch Situationen, in denen sie ihre Reputation nicht schädigen, zu ihrem persönlichen Vorteil aus. Jedoch selbst in völlig anonymen Situationen kann durch besondere Rahmenbedingungen, wie freiwilliges Teilnehmen an einer Gemeinschaft, Kooperation in einer sonst unkooperativen Situation gefördert werden.


Introduction

In this general introduction I will present the theoretical and methodological tools necessary to perform the experiments of the chapters I-V.

1 COOPERATION

Cooperative behavior is widely spread throughout the animal kingdom. Many species such as apes, lions, ancestral and modern humans are hunting in groups. The evolutionary puzzle of the emergence and sustainability of cooperative behavior within populations of selfish individuals is usually explained through kin selection (Hamilton, 1964b), group selection (Wilson and Sober, 1994), reciprocal altruism (Trivers, 1971) and mutualism (Brown, 1983; Connor, 1995). However, these models could not fully explain cooperation between non-related individuals, which modern human societies so strongly depend on the cooperation between individuals who are not related. Only recently it has been shown theoretically (Leimar and Hammerstein, 2001; Nowak and Sigmund, 1998b) that cooperation can evolve through indirect reciprocity, and empirically (Milinski et al., 2001; Wedekind and Milinski, 2000) that human subjects cooperate in indirect reciprocity situations (see also next section – 2 Game Theory) through reputation building. Punishment has been identified as another effective way of establishing cooperation between unrelated individuals (Boyd and Richerson, 1992; Fehr and Gächter, 2002; Gintis, 2000b; Sigmund et al., 2001). In addition several authors suggested that costly signaling could possibly provide an explanation for cooperation (Gintis et al., 2001; Leimar and Hammerstein, 2001; Roberts, 1998; Zahavi, 1995).

However, under some circumstances humans do not behave cooperatively. The anonymous use of limited resource with open access to everyone leads to an overuse of this resource, so that it cannot be sustained. Examples for such social dilemmas are unmanaged fish stock or hygiene in highly anonymous public places such as train stations. These situations are known as public goods situations (see also next section – 2 Game Theory). In empirical experiments the basic conditions of public goods situation are altered in order to find cooperative solutions, as I describe in the experiments of the chapters II - V.

2 GAME THEORY

Economist, social scientists and theorists have studied the problem of cooperative behavior in humans often using game theory as a tool for their investigations. J. v. Neumann and O. Morgenstern founded game theory in 1944 as a tool to predict the possible behavior of cybernetic systems (systems that have the ability to self regulate through feedback mechanisms when they are disturbed) in conflicting situations with the help of probability calculations. In this thesis I used two theoretical games. The first is the generally cooperative game of indirect reciprocity and the second is the usually uncooperative public goods game.

Indirect reciprocity

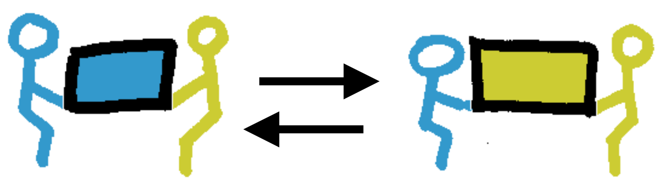


Fig. 2.1 Direct reciprocity: Blue helps Yellow and vice versa

Indirect reciprocity is an extension of direct reciprocity, where a donor gives help to a receiver and the receiver returns the help to the donor in the future (fig. 2.1). Nice indicators for the presence of direct reciprocity in human societies are sayings such as "if you scratch my back I will scratch yours". However, when direct reciprocity is excluded, so that the help cannot be returned by the receiver to the original donor, then it might be possible to help a third person, whereby the help may return indirectly to the original donor, which is then called indirect reciprocity (Fig 2.2). As an example Yellow gives help

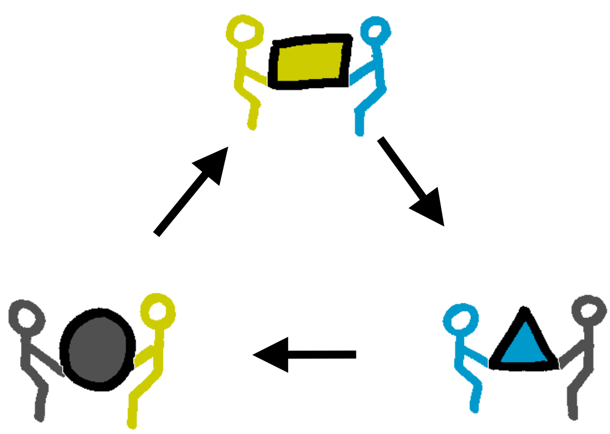


Fig. 2.2 Indirect reciprocity: Yellow helps Gray, Blue helps Yellow and Gray helps Blue.

to Gray who cannot return the help to Yellow, but Yellow is instead helped by Blue. Later Gray is able to help Blue, so that the help is returned indirectly. The biblical saying “give and you shall receive” is a nice example for indirect reciprocity.

It is known from theoretical and empirical work that humans cooperate in indirect reciprocity situations through reputation building. It has been shown theoretically and empirically that people who have given in the past are more likely to receive help in the future than people who refused to help (Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b; Wedekind and Milinski, 2000).

Public goods

The public goods games has first been described by Hardin as the “Tragedy of the Commons” (Hardin, 1968). The classic example of a public goods game consists of four players who can contribute money (e.g. one Euro) into a public pool (the public resource). The players can decide anonymously if they want to contribute. The content of the pool is then doubled and paid evenly to all players irrespectively of their contributions into the public pool. This situation is a social dilemma, because there is a conflict between the group interest and the individuals’ interest. The group stands best if all players contribute into the pool, then the group doubles its investment. However, the rational individual player should never contribute into the pool at all. The Euro paid into the public pool is doubled and then divided by four players. The investor receives only a return of 0.50 Cents and therefore only half of the investment.

Nevertheless Experiments with human subjects usually start highly cooperative, but when the game is played repeatedly the cooperation declines over time towards zero.

3 GENERAL APPROACH TO TESTING

All five experiments where carried out at the Universities of Bonn, Hamburg and Kiel, with first semester students from the biological faculties. Every experiment was done in cooperation with M. Milinski and H.-J. Krambeck. Furthermore T.C.M. Bakker was involved in the experiment described in chapter I. All volunteers from the biological first semester lectures, received a postal card with their appointed time and the room number, where the experiment would take place. Up to this point the students only knew that they would participate in an experiment, that they would play a simple group game and would receive money for doing so.

Fig. 3.1 Experimental setup: the experimenters' viewpoint with the controlling laptop computer in front.

When all volunteers had arrived they received an oral introduction. In this introduction we explained the use of the buttons for their decisions and that they would be completely anonymous throughout the game. The participants were also told that they should try to maximize their personal payoff throughout the game. Every participant was seated between separations (Fig. 3.1-3.3) and had a personal decision box with silent YES and NO buttons (Fig. 3.2). Additionally every participant received a pseudoname for the time of the game. To assure the anonymity, the participants had to choose their personal connecting cable from a bunch of similar cables, lying on the floor. The cables where disconnected after the game and intermixed. Finally, the participants where told that the computer program would start with a text introduction explaining the exact rules of the game (see Appendix). The complete game was displayed on a large screen in the front (Fig. 3.2, Fig. 3.3), so that complete information was accessible to all participants at all times. Throughout the experiment the experimenters were seated behind separations (Fig. 3.1), so that the players did not feel observed while playing the game. After each group had completed their game, the participants were asked to fill out a questionnaire to better understand the particular logic applied by the participants throughout the

Fig. 3.2 Experimental setup: view from behind the participants toward the large screen

experiments and to make sure they had understood the important rules of their specific game. The procedure of handing out the money that each participant had accumulated on her account had to insure the anonymity of the participants. For that reason we set up a kind of polling booth, where envelopes with the pseudonames of each participant containing the money that the computer program had calculated as the final amount, were spread out. Every participant had to step alone into the booth and take out all the money from the envelope with her pseudoname and return the empty envelope to the exact place where it had been before. In this way it was assured that the second participant would not know the pseudoname of the first participant and the last but one would not know the pseudoname of the last participant. Additionally, it was insured that the experimenters would not be able to link a pseudoname to a specific participant. Since it was also insured that the experimenters would not be able to trace any pseudoname back to a specific participant at any time, the experiment was conducted double blind. After all participants had received their money they were once more told not to talk about the experiment to others in order not to influence future participants. Furthermore, they were not allowed to tell their pseudoname to anyone in the future and were asked to forget this name if possible. All this was done so that the players could build up a new reputation exclusively by playing the game. Retaliation of uncooperative behavior or reward of cooperative behavior should be exclusively linked to the reputation that the players had achieved within the game.

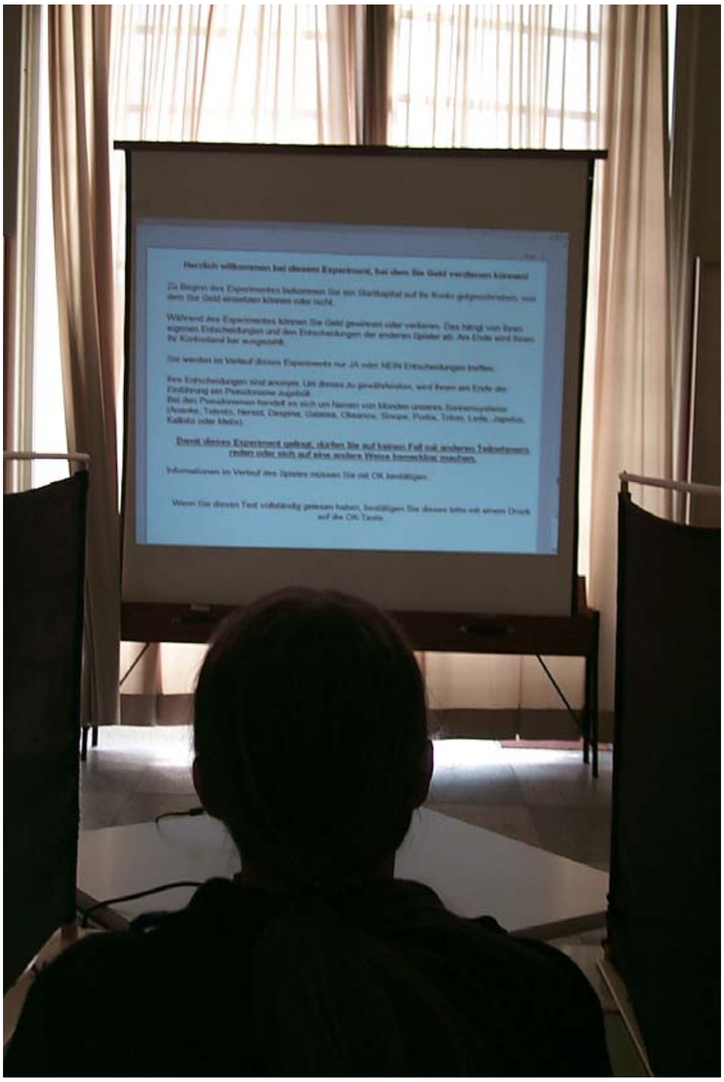


Fig. 3.3 Experimental setup: view from behind a single player

4 THESIS OUTLINE

This thesis is subdivided into 5 chapters. Each chapter is a single experiment investigating a new aspect of human cooperative behavior. This outline gives a short overview of the aim of each study and how it was tested.

Chapter I

In the first chapter I discuss the question why humans donate money to charity organizations. On first sight this seems to be an altruistic act. The direct fitness of the donator is reduced while the direct fitness of the recipient is increased. This is especially when the recipient is someone on a different continent (e.g. European charity organizations who collect money for aid in Africa). Reputation is known to be a strong force to establish cooperation in indirect reciprocity situations. The idea was that by making the donations to charity public they could possibly add to the individuals' reputation. Therefore we alternated in the experiment the donations to charity with indirect reciprocity rounds. In the indirect reciprocity rounds we displayed how the potential receiver had decided in past indirect reciprocity rounds and when she had donated to charity. After the experiment the participants voted whom of the other participants they would choose to be their group representative. We tested in this study whether donations to charity increased the likelihood of receiving from other players and in addition increase the likelihood of being voted the group representative.

Chapter II

In this chapter I investigate whether making decisions in public goods games reputation relevant can alter the usual uncooperative outcome of these games. In the experiment we alternated indirect reciprocity rounds with public goods rounds. In every indirect reciprocity round we displayed all past decisions of a potential receiver in the indirect reciprocity rounds and the public goods rounds. The prediction of this study was that the players would increase their donations into the public good in order to protect their good reputation for the indirect reciprocity games. Thereby the public good would be sustained and the group as a whole would benefit.

Chapter III

In the third chapter my aim was to investigate if it is possible to find a cooperative solution in the social dilemma of public goods even under complete anonymity and therefore the exclusion of reputation. A theoretical model (Hauert et al., 2002b) predicted that optional participation in the public goods should produce this effect. In this experiment each player has the choice between three strategies. First one could join the group and here could either cooperate (play as ”cooperator”) with a cost to pay for the cooperative act or to defect (play as “defector”) with no costs. The payoff depended in both cases on the number of players that had joined the group and how many of these players had decided to cooperate. The third strategy was not to participate in the group (play as “loner”) with a low but fixed payoff at the end of the round. Through this escape option from the social dilemma all three strategies were predicted to rise consecutively in an ongoing dynamic. This was predicted because whenever one strategy is the dominant choice within the group one of the other two strategies has a higher payoff, which we tested in a game with students.

Chapter IV

In the fourth chapter I looked at strategic behavior depending on the knowledge of being recognizable in a different social context or not being recognizable. Again as in chapter II the students played public goods and indirect reciprocity rounds. However in this experiment each participant had two different identities. The first name was as before used in the indirect reciprocity and in some public goods rounds. The decisions in the public goods rounds done with the first name would then be shown in all future indirect reciprocity rounds and were therefore linked to the individuals’ reputation. Whereas the second name was used exclusively in the remaining public goods rounds and the decisions were therefore not linked to the reputation in the indirect reciprocity rounds. We tested whether the participants used this information strategically. Predicting that they would reduce their cooperation with the second name, while increasing cooperation when making decisions with the first name in public goods rounds.

Chapter V

The purpose of my final chapter was to show the value of a good reputation inside and outside the own social group. While we could confirm in chapter I, II and IV that a good reputation within the own social group is highly valuable it was not known if this reputation is valuable outside the own social group. In this empirical study we let students play again public goods and indirect reciprocity games. The difference was that now there were two groups playing together all public goods rounds. For the indirect reciprocity rounds some of the players were exchanged between the two groups and these newly formed groups played all indirect reciprocity rounds together. Thereby the individual player had some players he was playing with in both situation and some players he would exclusively meet in either public goods or indirect reciprocity rounds. We tested here whether the participants treated the players who did not play in their own public goods group differently from the others in the indirect reciprocity rounds.

Chapter I

I DONORS TO CHARITY GAIN IN BOTH INDIRECT RECIPROCITY AND POLITICAL REPUTATION

Summary

Darwinian evolution can explain human cooperative behaviour among non-kin by either direct or indirect reciprocity. In the latter case one does not expect a return for an altruistic act from the recipient as with direct reciprocity, but from another member of the social group. However, the wide spread human behaviour of donating to poor people outside the social group, e.g. to charity organisations, that are unlikely to reciprocate indirectly and thus are equivalent to defectors in the game is still an evolutionary puzzle. We show here experimentally that donations made in public to a well-known relief organisation resulted both in increased income that the donors received from the members of their group and in enhanced political reputation: they were favoured to represent the interests of their group. Donations may thus function as honest signal for one's social reliability.

Introduction

It has been a longstanding evolutionary problem to understand how egoists can maximise their fitness by helping unrelated conspecifics (Nowak and Sigmund, 2000). Evolutionary theorists have developed mainly two concepts: direct and indirect reciprocity. In direct reciprocity (Axelrod, 1984; Axelrod and Hamilton, 1981; Milinski and Wedekind, 1998; Trivers, 1971) someone receives help and thereby gains more than the help costs the donor. If the help is reciprocated on the next occasion, each player has a net benefit. For indirect reciprocity (Alexander, 1987; Zahavi, 1991; Zahavi, 1995) support is given to individuals who have helped others. Both computer simulations and analytical models proved that indirect reciprocity can be evolutionarily stable (Leimar and Hammerstein, 2001; Lotem et al., 1999; Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b) and humans use it within their social group (Milinski et al., 2001; Milinski et al., 2002b, chapter II; Seinen and Schram, 2001; Wedekind and Milinski, 2000). Similarly, in the solidarity game players offer support to potential losers within the social group (Selten and Ockenfels, 1998). If, however, donations are given to non-members of the group, e.g. to charity organisations that help people in other countries, this kind of altruism may or may not be part of the indirect reciprocity game.

Since indirect reciprocity involves reputation and status (Alexander, 1987; Zahavi, 1991; Zahavi, 1995), a donation that is made in public may work as a conspicuous honest signal of a person's ability to participate in indirect reciprocity, or as Alexander (1986, p.100) put it: "In complex social systems with much reciprocity, being judged as attractive for reciprocal interactions may become an essential ingredient for success". Making donations in public to charity could in this way be explained by evolutionary theory. If acts of giving reveal important aspects of individual quality, there is the possibility that this information could be used also in other contexts (Leimar and Hammerstein, 2001; Zahavi, 1995), e.g. when deciding upon the delegation of powers to a person.

Methods

We tested these hypotheses with 72 students that participated in 12 groups of 7 subjects each in a computerised experiment. Each person had a starting account of DM 35 (about £17), and was anonymous with a pseudo-name (i.e., a name of a moon of our solar system). In each of the 16 rounds of the game each subject was assigned to be a potential receiver once and a potential donor twice, i.e. he/she was asked whether he/she would donate to a member of the group and thereafter to donate to charity. For example, a potential donor, say "Telesto", was asked whether he would give to "Galatea". Telesto would loose DM 2.50 from his account and Galatea would gain DM 4 on her account if Telesto decided YES. Telesto's decision (YES or NO) was displayed for 2s on a big screen which all participants could see all the time. Thereafter Telesto was asked whether he would give DM 2.50 to the relief organisation "UNICEF", which if YES would receive DM 4. This decision was also displayed for 2s. It was made clear that the money on UNICEF's account would be sent to UNICEF. Everybody was provided with information on whether everybody else, e.g. Galatea, had given in previous rounds (to other subjects and/or to UNICEF) when she/he had been in the role of the potential donor. The subjects knew that there would be no direct reciprocity. One student in each group had been secretly instructed by us to alternate YES and NO when asked to give to other players and, when asked to give UNICEF, to decide always YES ("YES-player") in 6 groups and always NO ("NO-player") in the other 6 groups. After the 16th round each subject was given a ballot and asked to elect a member of the group (pseudo-name) as a potential delegate in the students' council. This election had not been announced. Every subject received the money from his/her account anonymously after the experiment.

Results and Discussion

The subjects cooperated by indirect reciprocity, i.e. they received the more money the more they had given to others themselves (Fig. I.1). The amount of money given to others did not correlate significantly with the number of donations to charity (UNICEF) (r^2 = 0.006, df = 72, P >0.5, two-tailed, "YES" and "NO-players" excluded). However, those who had donated more to UNICEF received relatively more from other players (red dots in Fig. I.1), whereas those who had donated less to UNICEF received relatively less from other players (blue dots). To avoid pseudo-replication, the regression of donations to UNICEF on the residuals from the relationship "donations to others, donations received" was calculated for each group of 7 subjects separately, "YES" and "NO-players" excluded. The resulting 12 regression coefficients were on average (r = 0.36 + 0.11) significantly positive (Wilcoxon one-sample test against 0, z= 2.59, P <0.01, two-tailed). Donations to UNICEF thus paid off through indirect reciprocity. Similarly, the

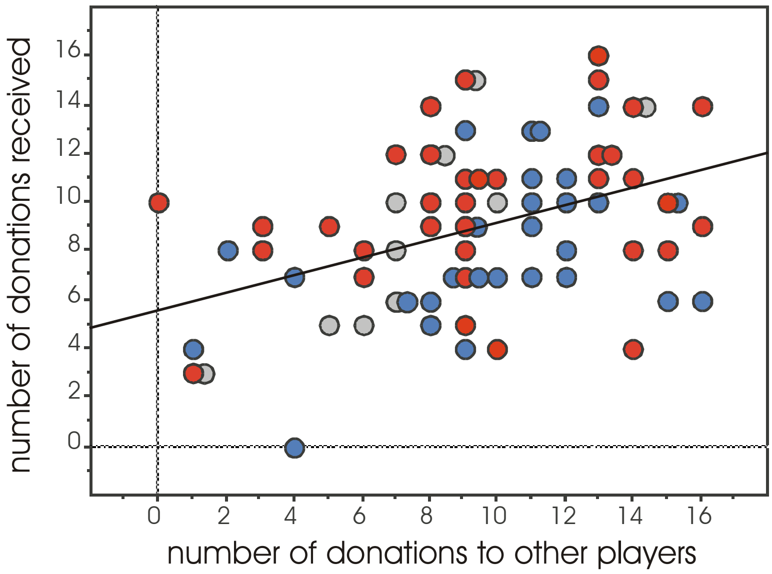


Fig. I.1 Human subjects received money indirectly related to the amount they gave to others (i.e. the more they gave to others the more they received; n =72, t = 3.71, P = 0.0004, two-tailed). The solid line depicts linear regression. Red circles are charitable donors (UNICEF) who gave more than the median, blue circles are donors who gave less than the median, and gray circles are median donors.

6 “UNICEF-NO-players” received significantly more NOs from their donors (52 + 11%) than did the 6 “YES-players” (30 + 3%, Mann-Whitney U-test, z = 1.93, P = 0.033, directed). UNICEF-YES-players thus received on average DM 12.80 more than UNICEF-NO-players; both types of pseudo-player did not differ in the amount they had donated to other players.

The sum of donations to UNICEF and to other players correlated positively with the number of votes that the subjects received in the election for the students’ council (Fig. I.2) (Spearman correlation, n = 84, z = 2.84, P<0.005, two-tailed). Since voting was by secret ballot and had not been announced we treat each subject as a statistical unit; the YES- and NO-players received votes but did not vote themselves. The number of donations to UNICEF (irrespective of the number of donations to other players) correlated positively with the number of votes received (regression of the number of votes received on the residuals from the relationship “donations to other players, donations to UNICEF”, n = 84, t = 2.60, P = 0.01, two-tailed). Similarly, the UNICEF-YES-players received 8 votes whereas the NO-players obtained only 3 votes. However, the number of donations to other players (irrespective of the number of donations to UNICEF) did not correlate significantly with the number of votes received (regression of the number of votes received on the residuals from the relationship “donations to UNICEF, donations to other players”, n = 84, t = 1.58, P = 0.12, two-tailed). This suggests that donations to charity have a stronger influence on political reputation than have donations to members of the group. This is corroborated by the finding that UNICEF-YES-players received on average about three times as many votes as did UNICEF-NO-players.

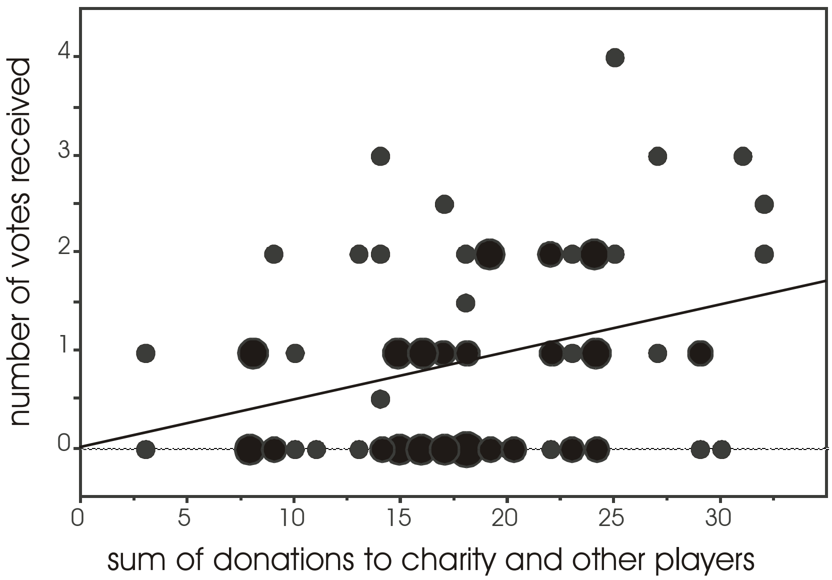


Fig. I.2 The number of votes the human subjects received in a staged poll for the students´ council was directly related to the amount they had donated to charity (UNICEF) and to the other players. The solid line depicts linear regression and circles of increasing size depict one to four subjects per data point.

However, each DM invested to charity by a UNICEF-YES-player gave a return of only DM 0.33 from indirect reciprocity. It might therefore be worth investing in both indirect reciprocity to gain primarily help from others and in charity to gain primarily another type of social reputation.

We have recently shown that the need to maintain reputation in the indirect reciprocity game can raise the level of contribution to a 'public good' considerably when both types of social dilemma are alternated (Milinski et al., 2002b, chapter II). It is possible that high reputation gained by donations to charity would further facilitate cooperation in a "tragedy of the commons". If it is made public that all participants of a public goods game did not give to charity, we would predict that the game starts already uncooperatively.

Our results show that donations to a relief organisation can pay off through both indirect reciprocity and increased reputation in another context, e.g. political eligibility. This result is compatible with Alexander's (1986) arguments: "Systems of indirect reciprocity, and therefore moral systems, are social systems structured around the importance of status... Status can be determined by physical prowess, as in those non-human (animal) dominance hierarchies in which coalitions are absent or (as in humans) by mental or social prowess. Mental and social prowess, in this sense, includes (as in moral systems) effectiveness and reliability in reciprocity and cooperation." It might (Nowak and Sigmund, 1998b) and, as we found, does pay to "advertise" cooperation. However, although donating to those who are in need might serve as an honest and efficient (because it is done in public) signal for one's reliability in reciprocity, this situation seems exploitable by defectors as has been masterly described by Sir Arthur Conan Doyle in one of his short stories (Conan Doyle, 1986). We proposed UNICEF for donations because its trustworthiness is beyond all doubt. Had we offered a less trustworthy organisation, donations to it might not have been as effective in raising a donor's status. This may depict a new dimension in the evolutionary arms race between cooperators and defectors in the tragedy of the commons (Hardin, 1968; Ostrom et al., 1999).

Acknowledgements

We thank the students from the University of Hamburg for participating in our experiment, J. Ganzhorn, M. Gewecke, H. Kiesewetter, J. Krink, J. Parzefall and H. Strutz for support.


Chapter II

II REPUTATION HELPS SOLVE THE 'TRAGEDY OF THE COMMONS'

The problem of sustaining a public resource that everybody is free to overuse - the 'tragedy of the commons' (Berkes et al., 1989; Dawes, 1980; Hardin, 1968; Hardin, 1998; Ledyard, 1995; Ostrom, 1999; Ostrom et al., 1999) - emerges in many social dilemmas, such as our inability to sustain the global climate. Public goods experiments (Ledyard, 1995), which are used to study this type of problem, usually confirm that the collective benefit will not be produced. Because individuals and countries often participate in several social games simultaneously, the interaction of these games may provide a sophisticated way by which to maintain the public resource. Indirect reciprocity (Alexander, 1987), 'give and you shall receive', is built on reputation and can sustain a high level of cooperation, as shown by game theorists (Lotem et al., 1999; Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b). Here we show, through alternating rounds of public goods and indirect reciprocity games, that the need to maintain reputation for indirect reciprocity maintains contributions to the public good at an unexpectedly high level. But if rounds of indirect reciprocation are not expected, then contributions to the public good drop quickly to zero. Alternating the games leads to higher profits for all players. As reputation may be a currency that is valid in many social games, our approach could be used to test social dilemmas for their solubility.

Since Hardin (Hardin, 1968) first described the 'tragedy of the commons', this type of social dilemma has been studied extensively by political and social scientists, economists and evolutionary theorists (Berkes et al., 1989; Dawes, 1980; Hardin, 1998; Ledyard, 1995; Ostrom, 1999; Ostrom et al., 1999). Many of the experiments that have been carried out are a variant of the standard design (Ledyard, 1995). In this model, four students seated at a table are each given an endowment of £5. They are then told that they can each choose to invest some or all of their £5 in a group project by putting, without discussion, an amount between £0 and £5 in an envelope.

The experimenter will collect the 'contributions', total them up, double the amount, and then divide this money among the group.

The economic/game-theory prediction is that no one will ever contribute anything because each £1 contributed yields only £0.50 to its contributor, no matter what the others do. This is a public goods problem because the group would be best off (taking home £10 each) if all contributed £5. But individual self-interest is at odds with group interest. Usually people cooperate more than is predicted by standard economic theory (Ledyard, 1995); however, observed cooperation is heterogeneous and declines over time (Fischbacher et al., 2001). It has been shown that direct punishment of non-cooperators can cause a rise in the level of the average contribution to the public good (Boyd and Richerson, 1992; Fehr and Gächter, 2000; Gintis, 2000a), and cooperators are even prepared to pay a cost for punishing ('altruistic punishment')(Fehr and Gächter, 2002).

We present an alternative way to maintain potentially a high level of contribution to the public good. It can be achieved through interaction with a second game that promises rewards for those with a good reputation in the public goods game. Theorists have shown that cooperation through indirect reciprocity can evolve (Lotem et al., 1999; Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b). For indirect reciprocity, individuals who have helped others are given support, whereby the supporter builds up reputation (Alexander, 1987; Zahavi, 1991) or a positive image score (Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b). Experimental studies have confirmed that human subjects preferentially help others who have a positive image score (Milinski et al., 2001; Seinen and Schram, 2001; Wedekind and Milinski, 2000). As players would risk their reputation if they would not cooperate in a public goods game that is alternated with the indirect reciprocity game, we predicted that alternating rounds of these two games would induce continuous cooperation in the public goods game, in contrast to a situation in which all public goods rounds were played first.

We tested these predictions with 114 first-year students who participated in 19 groups of 6 subjects each in a computerized experiment. The six subjects of each group could see a public screen on which instructions and the actual game was projected. They were told, first, that each person had a starting account of DM 20 (£10) and could gain or lose money dependent on his/her and the participants' decisions; second, that all decisions were anonymous and each player would be assigned a pseudoname (that is, a new identity) for the whole game; and last, that they would play in two different situations, an 'indirect reciprocity game' and a 'public goods game'.

Ten groups played one round of indirect reciprocity in which each subject was a potential donor once and a potential receiver once, and then one round of public goods. This alternating pattern was continued until round 16, thereafter four rounds of public goods were played. Every second group was told in round 17 that from then on only public goods rounds would follow until the end of the game. Nine other groups played eight rounds of public goods, followed by eight rounds of indirect reciprocity, followed by four rounds of public goods. Again, every second group was told in round 17 that from then on only public goods rounds would follow until the end of the game. In each round of an indirect reciprocity game, each potential receiver's history of giving both in the indirect reciprocity and the public goods game was displayed simultaneously for all players.

In groups that started with eight rounds of the public goods game initial cooperation declined as is usual in this type of game from round one to round eight (paired t-test between first and eighth round of public goods game, $n = 9$ groups, $t = 6.958$, $P < 0.0001$; Fig. II.1). During the subsequent eight rounds of pairwise indirect reciprocity, cooperation was instantaneously re-established (comparison between eighth round of public goods game and first round of indirect reciprocity: paired t-test, $n = 9$ groups, $t = 2.9$, $P < 0.02$; to avoid pseudoreplication we use each group of six subjects as our statistical unit throughout this paper; all probabilities are two-tailed).

But in groups that started with one round of indirect reciprocity, followed by one round of public goods, and so on until round 16, the initial high cooperation level of the public goods game did not decline during the eight rounds of the public goods game (comparison between the first and the eighth round of public goods game; paired t-test, $n = 10$ groups, $t = 0.897$, $P = 0.40$), and was on average considerably higher than the cooperation level of the nine groups that had started with eight rounds of public goods (unpaired t-test, d.f. = 17, $t = 4.83$, $P < 0.0002$; Fig. II.1). When public goods and indirect reciprocity rounds were alternated, the public goods game elicited significantly more cooperation than the indirect reciprocity game (comparison between average cooperation of eight rounds public goods and eight rounds indirect reciprocity; paired t-test, $n = 10$ groups, $t = 3.99$, $P < 0.004$).

The high cooperation level in the public goods game was probably maintained in the following way. Players might have withheld help in the pairwise indirect reciprocity game from players who had refused to give in the preceding public goods round. The probability of receiving 'no' in the indirect reciprocity game was significantly higher for players that had refused to

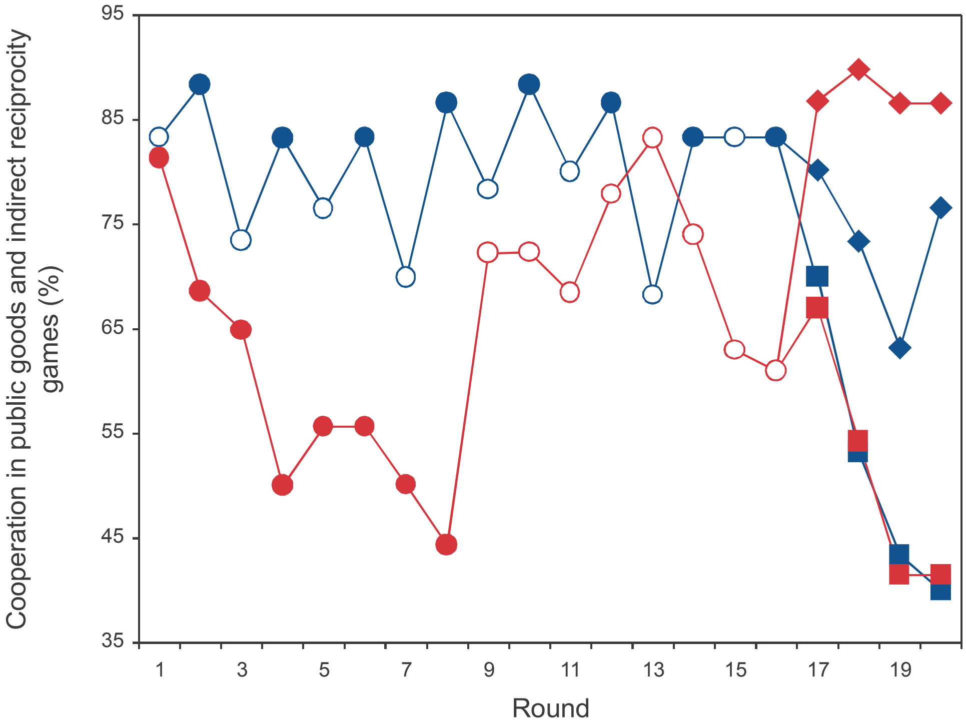


Fig. II.1 Percentage of cooperation ('yes') per group of six subjects in each round of the public goods game (filled symbols) and in each round of the indirect reciprocity game (open symbols). In one treatment, the groups alternated between rounds of indirect reciprocity and rounds of public goods until round 16 (blue); in the other treatment, groups started with eight consecutive rounds of the public goods game and continued with eight rounds of the indirect reciprocity game (red); in rounds 17-20, groups of both treatments played the public goods game, which was either announced, 'from now on only this type of game until the end' (squares), or not announced (diamonds).

give in the preceding public goods round than for those who had given (Fig. II.2a). Similarly, we found a positive correlation between the probability of receiving 'no' in the first round of the indirect reciprocity game and the rate of refused help during the block of eight rounds of the public goods game (mean Spearman's r per group = 0.49; s.e.m. = 0.15; Wilcoxon test against 0, n = 9 groups, z = -2.1, P < 0.04).

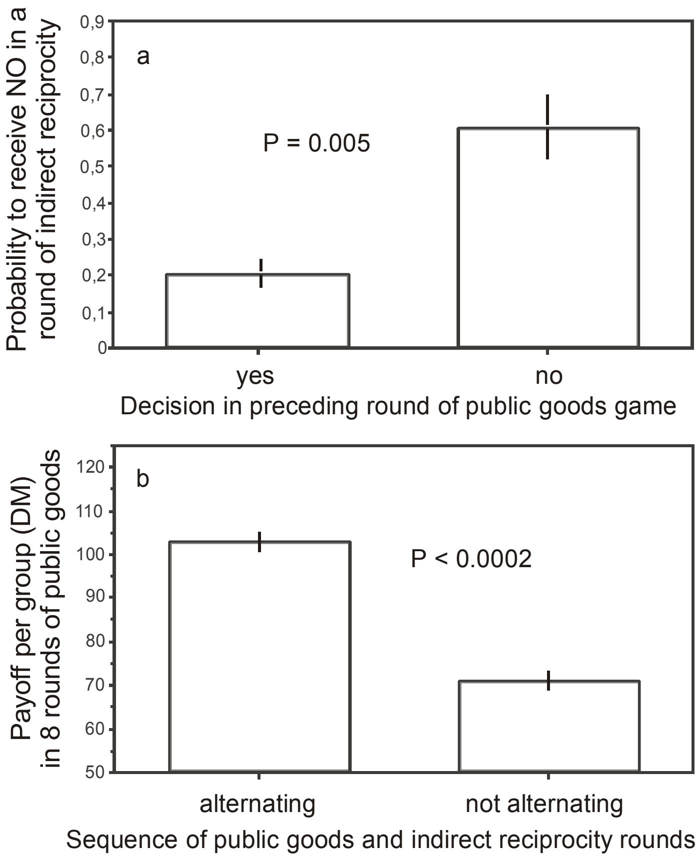


Fig. II.2 Consequences of cooperation in the public goods game. **a**, Probability of receiving 'no' in a round of the indirect reciprocity game depending on whether a subject had given either 'yes' or 'no' in the preceding round of the public goods game in the alternating treatment. The probability per group is shown (mean ± s.e.m) for both situations; all individual situations were taken to generate one mean value of either type for each group; paired t-test with arcsine-transformed data, n = 10 groups, t = 3.7, P = 0.005. **b**, Payoff (DM) per group (mean ± s.e.m.) in the first 8 rounds of the public goods game of all groups that either alternated the indirect reciprocity and the public goods game during the first 16 rounds or started with 8 consecutive rounds of the public goods game and thereafter played 8 rounds of indirect reciprocity; unpaired t-test, d.f. = 17, t = 4.83, P<0.0002.

The hypothesis that interaction with the indirect reciprocity game keeps up cooperation in the public goods game is directly tested by the four rounds of public goods that groups in both treatments played in rounds 17–20. Every second group was told in round 17 that from then on only public goods rounds would follow until the end of the game. In these groups cooperation declined during the four public goods rounds, whereas cooperation was maintained when the risk of further rounds of indirect reciprocity was not excluded (Fig. II.1; comparison of mean cooperation level during four rounds of public goods between five groups 'with' and five groups 'without announcement' after the alternating treatment, d.f. = 8, unpaired t-test, t = 4.456, P = 0.002, and between five groups 'with' and four groups 'without announcement' after the block treatment, d.f. = 7, unpaired t-test, t = 6.631, P = 0.0003). Thus, the pending risk of further rounds of indirect reciprocity prevented cooperation in the public goods game from declining at least over four consecutive rounds.

Obviously, refusing to give in the public goods game reduced the reputation of a player to a similar extent as if this person had refused to give in the indirect reciprocity game: his potential donor in the next round of indirect reciprocity just followed the rules for indirect reciprocity and refused to give to someone with a low image score. This is different from punishing because it does not need any special punishing rule or motivation, and the potential donor actually saves money by refusing to give. A recent theoretical analysis(Sigmund et al., 2001) suggests that reputation is essential for fostering social behaviour among selfish agents, which is confirmed experimentally here. The inclusion of reputation effects in the corresponding dynamical models leads to the evolution of economically productive behaviour, with agents contributing to the public good and either punishing those who do not or rewarding those who do (Sigmund et al., 2001). Providing help in the indirect reciprocity game is a form of reward.

Cooperation in the public goods game paid off. Groups that alternated rounds of indirect reciprocity and public goods games, and thus were more cooperative in the public goods game, earned significantly more money during the first eight rounds of the public goods game than did groups that played the two games in blocks of eight rounds each (Fig. II.2b). This shows that the 'tragedy of the commons' was no longer a tragedy; instead, the commons became productive and could be harvested. Two people usually interact in more than one situation, therefore their actions in one context may influence actions in another (Coleman, 1990). Many social dilemmas are a type of public goods game (Ostrom et al., 1999), others have been identified as a type of indirect reciprocity game (Ledyard, 1995). It therefore seems likely that the kind of interaction that we have staged experimentally occurs naturally in our society. There might be hidden social dilemmas that would show up only if the interaction with another game were removed.

Methods

Indirect reciprocity game

Players were anonymous; each subject was assigned a pseudoname by the computer for the whole session of 20 rounds so that at any time, players could make their decisions contingent on the history of the game up to that time; each player knew his/her name but did not know who had been assigned the other names; the subjects were separated by opaque partitions and communicated their decisions with silent (piezo) switches; they knew that they would obtain their money after the game in a way that did not disclose their anonymity.

For the 'indirect reciprocity game'20, each person was assigned repeatedly as either a potential donor or a potential receiver. For example, a potential donor, say 'Telesto', was asked on the public screen whether he would give to 'Galatea'. Telesto would loose DM 2.50 from his account and Galatea would gain DM 4 on her account if Telesto decided 'yes'. Telesto's decision (yes or no) was displayed for 2 s on the public screen. Everybody knew about the contributions of all players, for example, whether Galatea had given in previous rounds when he/she had been playing as the potential donor. The subjects also knew that there would be no direct reciprocity; if A was the potential donor of B, B would never be the potential donor of A. In each round of the indirect reciprocity game, each of the six players was once a potential donor and once a potential receiver.

Public goods game

For the 'public goods game'4, all six players were asked simultaneously whether they would contribute DM 2.50 to the public pool, the contents of which would then be doubled and redistributed evenly among all players irrespective of whether they had contributed. After all players had decided, each player's decision (yes or no), his/her contribution (that is, DM 2.50 or 0), and his/her gain (for example, DM 4.17 if all but one had contributed), was displayed below the pseudonames for 20 s.

Acknowledgements

We thank students at Hamburg University for their participation; E. Heinz and W. Kiesewetter for support; E. Fehr and K. Sigmund for comments; and the Max Planck Institute of Meteorology at Hamburg for hospitality.

Chapter III

III VOLUNTEERING LEADS TO ROCK-PAPER-SCISSORS DYNAMICS IN A PUBLIC GOODS GAME

Collective efforts are a trademark of both insect and human societies (Trivers, 1985). They are achieved through relatedness in the former (Hamilton, 1964a) and through largely unknown mechanisms in the latter. The problem of achieving cooperation among non-kin has been described as the ‘tragedy of the commons’ prophesying the inescapable collapse of many human enterprises (Hardin, 1968; Hardin, 1998). In public goods experiments, initial cooperation usually drops quickly to almost zero (Ledyard, 1995). It can be maintained by either the opportunity to punish defectors (Fehr and Gächter, 2002) or the need to maintain good reputation (Milinski et al., 2002b, chapter II). Both scenarios require defectors being identified. Recently, theorists proposed a simple but effective mechanism operating under full anonymity. With optional participation in the public goods game “loners”, i.e. those players who do not join the group, defectors and cooperators will coexist through a rock-paper-scissors dynamics (Hauert et al., 2002a; Hauert et al., 2002b). Here we show experimentally that volunteering easily generates this dynamics in public goods games and that manipulating initial conditions can produce each predicted direction – if, by manipulating displayed decisions, defectors are pretended to have the highest frequency, loners soon become most frequent, as do cooperators after loners and defectors after cooperators, respectively. On average cooperation is perpetuated at a substantial level.

Clean air to sustain the global climate and clean public toilets are examples of public resources which everybody is free to overuse. The social dilemma of public goods situations is that although a group of cooperators is always better off than a group of defectors, defectors exploit cooperators within groups. Since the late 1970’s economists, social scientists and evolutionary biologists have used the public goods game as a paradigm to study the problem of maintaining cooperation within a group of unrelated individuals (Berkes et al., 1989; Colman, 1995; Gintis, 2000a; Ostrom, 1999; Sugden, 1986): for example, six players are asked to contribute money to a public pool; the money in the pool is, e.g., multiplied by 3.6 and equally distributed among the players irrespective of whether they contributed. The optimum outcome for the group is achieved if everybody cooperates. However, since each euro paid into the pool yields only a return of 60 cents for the contributor, i.e. a net deficit of 0.40 €, no matter how the other players decide, the selfish decision is never to contribute to the pool. Recent studies have identified punishment (Boyd and Richerson, 1992; Fehr and Gächter, 2000; Fehr and Gächter, 2002; Gintis, 2000b), also combined with fairness (Fehr and Rockenbach, 2003), and reputation through interaction with other social behavior(Milinski et al., 2002b, chapter II) as mechanisms that can effectively maintain cooperation in public goods experiments.

In their model Hauert et al. (Hauert et al., 2002a; Hauert et al., 2002b) consider a large population with three types of players: cooperators, defectors and loners. From time to time, sample groups of N players are randomly chosen and offered to participate in a single public goods game. Players can either refuse to participate, and then receive a small fixed payoff, or join the public goods game. In the latter case they either defect or cooperate. Their strategies are specified beforehand, and do not depend on the composition of the group. An ongoing oscillation of the three strategies is predicted because each strategy when most frequent can be beaten by one of the others. Defectors can exploit a large group of cooperators, whereas loners have the highest profit when defectors are frequent. When loners are most frequent the public group size is reduced, which invites cooperation because in small groups the game is no longer a dilemma (Boyd and Richerson, 1988; Dawes, 1980; Schelling, 1973). For example, if the group consists of only three players, each euro paid into the public pool yields a return of 1.20 € for the contributor, i.e. a net gain of 0.20 €. It is not just the fact that volunteering is possible that induces cooperation, but rather that volunteering reduces public goods groups to small sizes for which the individual cost to benefit ratio becomes more favourable. Furthermore, even though defectors are still better off than cooperators within each group, cooperators do better when averaged over small groups, according to Simpson's paradox(Sober and Wilson, 1999): e.g., a group of three players can consist of either 3 cooperators, 2 cooperators and 1 defector, 1 cooperator and 2 defectors, or 3 defectors. Cooperators receive on average a net gain of 1.8 €, defectors only 0.8 €. Circumstantial evidence for the "small group

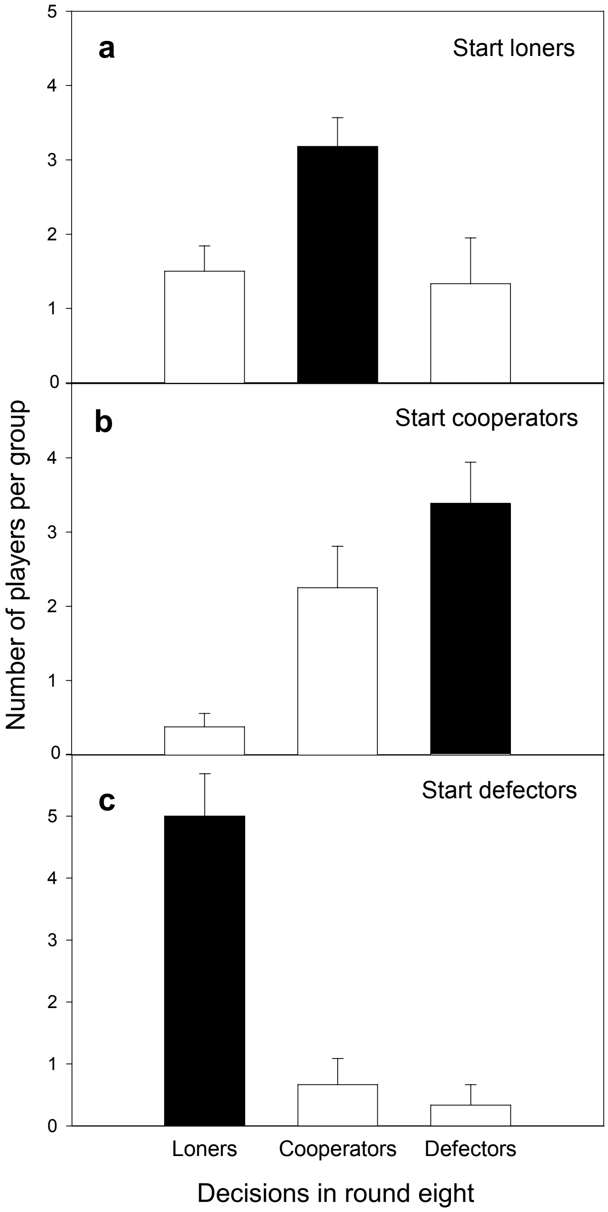


Fig. III.1 Decisions in round eight after the staged standstill with one single strategy that was most frequent during the first seven rounds, the bars of the predicted strategy to be most frequent are black. The average frequency of chosen strategies per group of 14 players is shown (mean ± SE). **a**, Start loner; n=6 groups of 14 players each with simulated prevalence of loners during round 1 to 7. **b**, Start cooperators; n=8 groups of 14 players each with simulated prevalence of cooperators during round 1 to 7. **c**, Start defectors; n=6 groups of 14 players each with simulated prevalence of defectors during round 1 to 7.

advantage" is potentially provided by fish that leave their shoal and take the risk to inspect a predator from a short distance: minnows, *Phoxinus phoxinus*, inspect a pike, *Esox lucius*, very often in small groups.(Magurran and Pitcher, 1987) So after loners cooperators will be most frequent for a while before defectors will take over again (Hauert et al., 2002a; Hauert et al., 2002b). Hence, volunteering relaxes the social dilemma: instead of defectors winning the world, coexistence among cooperators, defectors and loners is expected (Michor and Nowak, 2002).

We tested these predictions with 280 first semester biology students in twenty groups of 14 students each that played the optional public goods game for 57 consecutive rounds. The students observed the introduction and the complete game on a public screen. They were told that they had a starting account of 10 € and would make their decisions anonymously. For each round six players were randomly selected from a "population" to decide first whether they join the public goods group and thereafter, if they chose to join, whether they like to contribute to the public pool. During the first seven rounds we manipulated the displayed decisions in such a way that defector, cooperator or loner was pretended to be the most frequent strategy of the population. This manipulation was necessary to test the three possible predictions of the model experimentally. Without this manipulation our results would be only descriptive. In the eighth round we expect that being loner (after staged defector), defector (after staged cooperator) or cooperator (after staged loner), respectively, would be the most frequent strategy according to the players' real decisions. Thereafter the game proceeded with unmanipulated display to test whether oscillations of the three strategies occur and if so whether they occur predominantly in the predicted sequence during 50 rounds.

Following the manipulated start during the first seven rounds we found that the predicted strategy was the most frequent strategy after all three starting scenarios in round eight (Fig. III.1 a-c) ($P<0.004$, $n=20$ groups, sign test, two-tailed). We use each group of 14 players as statistical unit.

During the following 50 rounds we determined for each group of 14 players the number of cases where two conditions were met: one strategy was most frequent and one of the other strategies was the most frequent strategy in the following round. We compared all cases where the predicted strategy became most frequent with all cases where an unpredicted strategy became most frequent. So we uncovered switches of the most frequent strategy between rounds and checked whether their direction was as predicted. The predicted strategy became most frequent significantly more often than the alternative strategy (Fig. III.2) ($P<0.001$, $n=20$, paired t-test, $t=6.588$, two-tailed).

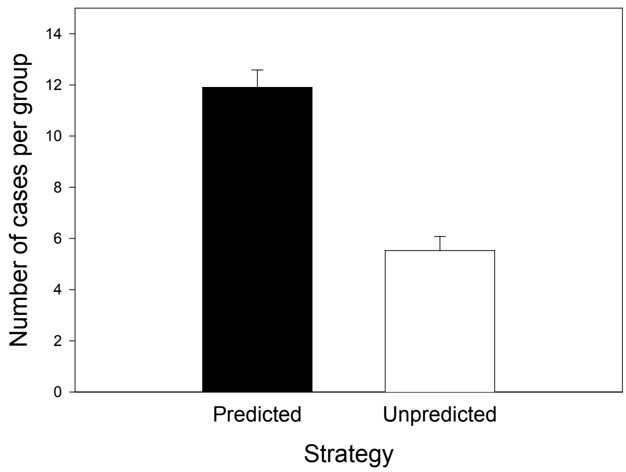


Fig. III.2 The predicted prevalence switch occurred more frequently than the unpredicted prevailence switch during the 50 rounds that followed the seven manipulated start rounds. Columns show mean ± SE per group of 14 players.

Although the above analysis provides a formal proof of the predicted oscillations we made an example of these oscillations visible in Fig. 3. We synchronized the 20 groups during the 50 not manipulated rounds by selecting similar starting points in each group, because the cycles were not expected to have the same duration in each group. For example if we select that round from each group which has the highest proportion of loners, we would expect that cooperators would be most frequent next, followed by defectors in all 20 groups. The same procedure was used to find such starting points for cooperators and for defectors. From several maxima of a strategy we defined the first as the starting point. Thereafter we averaged all 20 starting point rounds and each of the following 9 rounds over all groups, for loners (Fig. III.3a), cooperators (Fig. III.3b) and defectors (Fig. III.3c), respectively, as starting points. The oscillations can be observed in all three cases, although the groups became increasingly asynchronous during the ten rounds. As the model predicts, after loners have the highest frequency cooperators follow to become most frequent, thereafter defectors and again loners (Fig. III.3a). After a prevalence of cooperators defectors become most frequent followed by loners and again cooperators (Fig. III.3b). Fig. III.3c shows that the prevalence of defectors is followed as predicted

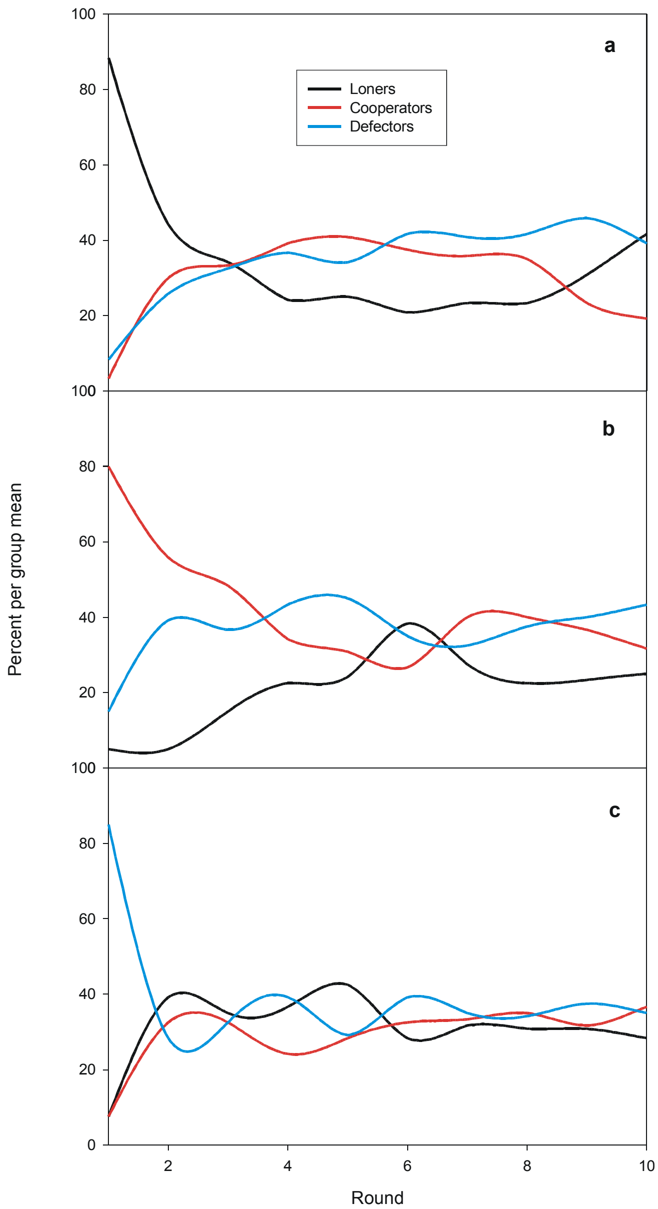


Fig. III.3 Average frequencies of the three strategies over a period of ten rounds after synchronizing the 20 groups as follows. The starting round of the ten rounds was defined for each group as the round when one of the strategies reached its maximum proportion throughout the game (rounds 8-57) for the first time. All 20 starting rounds and each of the following 9 rounds were averaged over all groups, for loners **a**, cooperators **b** and defectors **c**, respectively. Sequence of predicted most frequent strategies according to the rock-paper-scissors dynamics: loners, cooperators, defectors, loners, and so on.

by an increase of loners that is closely trailed by increasing numbers of cooperators followed again by defectors and thereafter by loners.

The consequences of the oscillation of the strategies should be an always recurring rise of each of the three strategies and thus a fairly cooperative outcome of the game after initial perturbations. In the last thirty rounds (21-50) the frequencies of the three strategies appeared on average rather stable (round 21-35: 32.22+1.0% loner, 30.11+0.9% cooperators, 37.67+1.0 % defectors and round 36-50: 32.39+1.4% loner, 29.06+1.3% cooperators, 38.56+1.3% defectors). According to the model we expect at least 42% loners and 58% should choose to join the public goods group(Hauert et al., 2002a). Only 33±2.5% (mean±SE) chose the loner option, which is significantly less than expected (P=0.003, n=20, Wilcoxon signed rank test, Z=2.95, two-tailed). Of the players joining the public goods group 38% are expected to cooperate and 62% to defect, respectively(Hauert et al., 2002a). We found, as expected, more defectors (56.51±1.7%) than cooperators (43.48±1.7%, P=0.004, n=20, Wilcoxon signed rank test, Z=2.91, two-tailed). Although these numbers are close to the expected ones, the percentage of cooperators was significantly higher than predicted (P=0.011, n=20, Wilcoxon signed rank test, Z=2.54, two-tailed).

In the long run, i.e. averaging over many cycles, the net payoff of both defectors and cooperators should be same as the loners' payoff, i.e. 1.25 €. We found that defectors earned slightly but significantly more than expected, i.e. 1.46±0.04 € (P<0.001, n=20, Z=3.36, Wilcoxon signed rank test, two-tailed). Cooperators had a payoff that did not significantly differ from the expected one, i.e., 1.32±0.09 € (P=0.43, n=20, Z=0.78, Wilcoxon signed rank test, two-tailed). Defectors probably profited because they were less frequent than expected at the equilibrium.

We found that volunteering, i.e. the option to choose between joining the public goods group and the loner strategy, indeed protected cooperation in the public goods game by inducing small group sizes. There was on average a rather stable frequency of cooperators that was higher than what is usually found in public goods games after several rounds (Fehr and Gächter, 2000; Ledyard, 1995). As predicted by the model (Hauert et al., 2002a; Hauert et al., 2002b) the dynamics of the games displayed oscillations of the rock-paper-scissors succession of cooperators, defectors, and loners even though our players were a bit less risk averse than expected: only about a third chose the loner option.

Volunteering is a mechanism, which potentially sustains cooperation in various species. Like some large predatory animals, also ancestral humans acted as groups when hunting large prey, e.g. mammoths, and went out solitarily for small prey, e.g. antelopes (Ridley, 1996). Thus, volunteering was possible and might have supported cooperation in addition to potential relatedness by reducing the public (hunting) group size. Obviously, we are not free to decide whether we stop sharing the global climate with others, but there are many other human social dilemmas in which volunteering is possible. Volunteering does not produce overwhelming cooperation but might help avoiding the fate of mutual defection in many human collective enterprises and thus pave the way for other mechanisms of cooperation to take over. For example, direct (Axelrod and Hamilton, 1981) or indirect reciprocity (Bolton et al., 2001; Nowak and Sigmund, 1998b; Seinen and Schram, 2001; Wedekind and Milinski, 2000) may be catalysed when the population happens to be in a cooperator period of the rock-paper-scissors dynamic and anonymity is relaxed after repeated interactions. Loners, although unsocial by definition, help cooperators to become most frequent and thus to escape the social dilemma.

Methods

The total of 280 human subjects of the universities of Bonn, Hamburg and Kiel played a public goods game with optional participation that lasted for 57 rounds. The students were completely anonymous, sat between partitions, saw the introduction to the game including one example round and the complete game on a large screen. They did not know the total number of rounds. They interacted via a computer program using silent “yes” and “no” switches.

For each round the computer program randomly selected six of the 14 students. Each student had played almost the same number of rounds at the end of the game. Since the expected cycles are predicted to become smoother with increasing population size (Hauert et al., 2002a), we mimicked a larger population. The students were told that there was a pool of additional players in the form of strategies recorded from earlier sessions and that the program would sometimes choose “players” from this pool. A light at each person’s desk signalled who was to decide. Each of the six players had to decide first whether to play the loner strategy thereby obtaining a fixed payoff (1.25 €) or to join the public goods group with a second decision to make. The minimum public group size was two players. If only one player decided to play in the public goods group she knew that she would automatically become also a loner. If the public goods group size was either two or larger, the players that had chosen to play in the public goods group had to decide whether they would contribute 1.25 € or nothing to the public pool. At this time they did not yet know how many subjects had decided to play in the public goods group. After all players of the public goods group made their final decision the content of the pool was multiplied by 3.6 and divided evenly among the players that had joined the public goods group irrespective of their actual contribution. With an interest rate of 3.6 the model system has a fixed point, which refers to substantial proportions of cooperators, defectors and loners. The dynamic then predicts periodic cycles of all three strategies around these levels, this requires an interest rate larger than 2. Only now the decisions of all players were displayed simultaneously on the screen that all 14 subjects could see: i.e., the numbers of loners and public good group players, their payoffs and their eventual costs, (e.g., one player was a loner and obtained 1.25 € without cost, five had chosen to join the public goods group, of which three were defectors who received a payoff of 1.80 € from the pool without costs and two were cooperators who also received 1.80 € from the pool, but they had costs of 1.25 € each). It never happened that one subject had to play loner because he had no money left.

During the first seven rounds the display was manipulated such that the players were lead to believe that they were in a group that played a high percentage of only one strategy. In six groups loners appeared to be most frequent; there were eight groups with cooperators and six with defectors as the apparent most frequent strategy. The players could make decisions, which were, however, not displayed. Instead, six predetermined decisions with corresponding payoffs and eventual costs were shown. Each of the three possible real decisions of a player (i.e. loner, cooperator and defector) was included at least once to ensure that each player would find his actual decision on the screen and nobody would doubt that the displayed decisions were real. The maximum number of defectors or cooperators displayed on the screen was 4 players per round; e.g. 4 defectors, 1 cooperator, 1 loner. In the case of loners prevailing the first seven rounds, it was possible to show up to 100% loners in one round, because each player who decided to join the public goods group would believe that he/she was the only one with this decision and thus became a loner. In order to have some variation we chose the average percentage of loners somewhat lower during the seven rounds. On average there were 79% loners in the staged loner groups, 61% cooperators in the staged cooperator groups and 64% defectors in the staged defector groups. Starting with round eight there was no manipulation of the display for 50 consecutive rounds. The students did not know the total number of rounds to be played.

Acknowledgements

We thank the students at the universities of Bonn, Kiel and Hamburg for their participation; T. Bakker, H. Brendelberger, E. Heinz, K.-P. Sauer and M. Zbinden for support, C. Hauert for calculating parameters, and the Max-Planck-Institute of Meteorology for hospitality.

Chapter VI

IV A HUMAN COOPERATION STRATEGY THAT IS CONDITIONAL ON BEING RECOGNIZED IN OTHER SITUATIONS

Abstract

Although collective efforts are common in both animal and human societies, many human and probably animal social dilemmas have no obvious cooperative solution, which is a challenge for evolutionary biologists. In public goods games, i.e. the experimental paradigm for studying the sustainability of a public resource with human subjects, initial cooperation usually declines quickly. Recently it has been shown that the interaction with another social game in which good reputation attracts help, can maintain a high level of cooperation in the public goods game. Here we show experimentally that humans use different strategies in the public goods game conditional on whether the player knows that his decisions will be either known or unknown in another social game. The knowledge of being recognized as the same individual in both scenarios motivates players to invest in their reputation and thus in sustaining the public resource. However, cooperation declines immediately when individual identities switch from being recognizable to being unrecognizable between the two interacting games.

1. Introduction

Cooperative behavior such as hunting in groups is known from several species, e.g. chimpanzees, lions, archaic and modern humans. Many aspects of present human societies depend on cooperation in order to function properly. The evolution of cooperative behavior within populations of selfish individuals is usually explained through either kin selection (Hamilton, 1964b), mutualism or reciprocal altruism(Axelrod and Hamilton, 1981; Trivers, 1971). Recently theorists (Leimar and Hammerstein, 2001; Lotem et al., 1999; Mohtashemi and Mui, 2003; Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b) have shown that cooperation can evolve also through indirect reciprocity (Alexander, 1987), “give and you shall receive”. By helping others, who do not have the possibility to return the help to the donor in the future, people build up good reputation or a positive image score, whereas refusing to help damages the reputation. Empirical studies confirmed that human subjects who have been helpful in the past are more likely to receive help from others through indirect reciprocity (Bolton et al., 2001; Milinski et al., 2001; Milinski et al., 2002a, chapter I; Seinen and Schram, 2001; Wedekind and Milinski, 2000).

This is, however, different in “public goods situations” which are typical social dilemmas where cooperation declines (Fischbacher et al., 2001). Social scientists, economists and evolutionary theorists have studied public goods situations extensively (Berkes et al., 1989; Dawes, 1980; Hardin, 1998; Ledyard, 1995; Ostrom, 1999) since Hardin first described this type of social dilemma as the ‘tragedy of the commons’ (Hardin, 1968). The classic public goods game consists of four players, who are given the opportunity to contribute money into a public pool. The content of the pool is doubled, divided by the number of players and evenly paid to all players, irrespective of their contributions. The social dilemma lies in the conflict between the group and the individual’s interest. The group does best when all players cooperate. However, a rational individual should never contribute anything, because each money unit paid into the pool yields only a return of a half-unit to the contributor. Thus, a limited public resource, which everyone is free to use, e.g. the global climate, unmanaged fish stock in common fishing grounds, or hygiene in highly anonymous public places such as train stations, is usually not sustained. There are certainly numerous potential scenarios where microorganisms (Rainey and Rainey, 2003; Velicer and Yu, 2003) or animals are found in public goods situations, e.g., when several individual parasites grow in an intermediate host, this “public resource” would not be sustained if each parasite would take as much energy from its host as if it would be alone (Christen and Milinski, 2003; Parker et al., 2003). Nonetheless there are several examples from human societies where the social dilemma has been successfully avoided by mechanisms such as control of access to the public good by the local community (Berkes et al., 1989).

Recently it has been shown that potential punishment of uncooperative group members (Boyd and Richerson, 1992; Fehr and Gächter, 2002; Gintis, 2000b; Sigmund et al., 2001), costly signaling with altruistic acts (Gintis et al., 2001), voluntary participation in the public goods game (Hauert et al., 2002b; Semmann et al., 2003, chapter III) and the interaction with indirect reciprocity situations can help solving the “tragedy of the commons” (Milinski et al., 2002b, chapter II). In the latter study groups of human volunteers played public goods games alternated with indirect reciprocity games. This alternation produced a high level of cooperation in the public goods games. Cooperation was maintained throughout the experiment, except when groups were informed that the last rounds would consist only of public goods games. The decline of cooperation in these groups suggests that the decisions made in the public goods games were relevant for the player’s reputation in the indirect reciprocity games. An alternative explanation for the observed break down of cooperation is the following: the interaction between the two games leading to potential information overload because of the limited channel capacity of the brain (Broadbent, 1965; Milinski, 1990) could have resulted in cooperative decisions. Removing the interaction between the two games would have removed this overload thereby allowing for uncooperative decisions again.

In the present experiment we did not remove the interaction between the two games. Instead we allowed for reputation transfer from the public goods game to the indirect reciprocity game in one treatment but did not allow for this transfer in the other treatment. If we find a higher level of cooperation in the public goods game when we allow for reputation transfer, this potential for reputation transfer must have caused the rise of investment in the public good. We achieved the manipulation of the reputation transfer by providing the subjects each with two different new identities, i.e. two pseudonyms. Each participant received two names of moons of our solar system, e.g. Telesto, Kalisto, ect.. One name was used only in public goods rounds whereas the other name was used in rounds of both games. With this procedure we test whether human subjects make strategic use of their knowledge of being recognized or not recognized as the same individual in both scenarios. We test whether this knowledge motivates players to invest in their reputation and thus in sustaining the public resource.

Humans may meet their neighbors repeatedly in various social games and should expect that their reputation would be transferred among games. However, when visiting other neighborhoods or other villages it may be rewarding to be uncooperative in a public goods game unless gossip finds the way home.

2. Method

We conducted our experiment with 120 students of the Universities of Bonn, Hamburg and Kiel. Each group consisted of 6 students, who were anonymous with respect to their real identity but were provided with two new identities, i.e. pseudonyms, under which they were recognized throughout the game. Thus, during the game the players learned about the decisions of other players only under these pseudonyms. Separated from each other, all players could observe the complete history of the game on a large screen and communicate their decisions through silent "yes" and "no" buttons at their desks. An oral introduction informed about the assignment of pseudonyms, the use of the silent switches and the procedure of the introductory part of the computer program (Milinski et al., 2001), which explained by means of both text and example rounds the rules of the game and provided each student with a starting account of 10 € and two different pseudonyms. The participants were informed that nobody including the experimenters could find out which pseudonym belonged to which real name. To assure the participants of this fact they were asked to choose a cable from a knotted bunch of identical cables. The chosen cable was then connected to the decision box at the participant's desk. After the last round the cables were disconnected and intermixed in front of the participants. This procedure was necessary to perform the experiment double blind, to avoid a rise in cooperation simply due to the fact that the participants did not fully believe in their anonymity (Hoffman et al., 1996).

The students played a mixture of public goods (PG) rounds during which all 6 players made their choices simultaneously and indirect reciprocity (IR) rounds with pair wise interactions. In each of the PG rounds the players could contribute 1.25 € from their account into the public pool. The content of the pool was then doubled and evenly distributed among all players irrespective of whether they had contributed. All the decisions, costs of the decisions and the money paid to the players from the public pool were simultaneously displayed after the last player had made her decision. Every IR round consisted of two interactions for each of the 6 players, once as the potential donor and once as the potential receiver. The subjects knew that the same two players could meet again in the same roles but never in alternated roles, so direct reciprocity was excluded. If a potential donor decided to donate, 1.25 € were taken from the player's account and 2.00 € were credited to the receiver's account. Since the value of the help received should be higher than the value of the costs for the donor (Nowak and Sigmund, 1998a), 0.75 € was added to the amount given. Before a potential donor would make her decision, some information about the potential receivers behavior as a donor in earlier rounds was displayed.

The first pseudonym ("reputation name") was used in all IR rounds. During each IR round the past decisions of the potential receiver of all the rounds where the reputation name had been used were displayed. Therefore all decisions made in IR rounds were shown in the future IR rounds. The reputation name was also used in some of the PG rounds, so the participants new that a decision made with the reputation name in a PG round would also be displayed in all future IR rounds. In the remaining PG rounds the second pseudonym ("reputation-free name") was used. The players were informed that the decisions of these rounds would never be displayed in any future rounds.

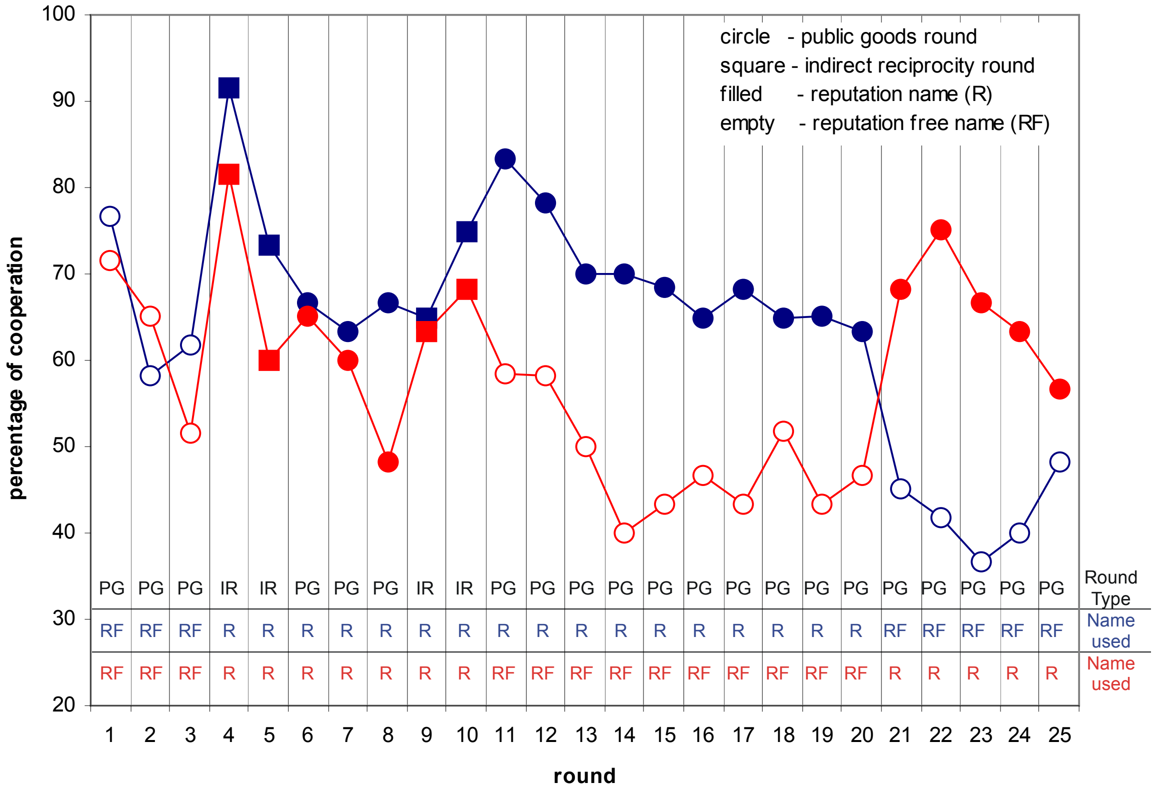


Fig. IV.1 For the public goods rounds (circle symbols) and indirect reciprocity rounds (square symbols) the group mean yes per round for both treatments are shown. In treatment one (blue) the groups played PG rounds, from round 11 to round 20 with their reputation name (R) (filled symbols) and from round 21 to 25 with their reputation-free name (RF). In treatment two (red) the groups played PG rounds, from round 11 to round 20 with their reputation-free name and from round 21 to 25 with their reputation name. The period from round 1 to 10 was in both treatments identical (three PG rounds played with the reputation-free name, two IR rounds with the reputation name, three public goods rounds with the reputation name and two IR rounds with the reputation name).

Each group started with 3 PG rounds using the reputation-free name, followed by 2 IR rounds, where the decisions of the first 3 rounds were not displayed (see fig. IV.1). Rounds 6 to 8 were PG rounds with the reputation name. Rounds 9 and 10 were IR rounds, where all previous decisions made with the reputation name were displayed. This introductory part was the same for both treatments to help the students to become accustomed to the procedure. The last 15 rounds (11 to 25) were the actual test in which the 2 treatments differed in order to control for sequence effects. 10 groups played 10 PG rounds with the players' reputation names, followed by 5 PG rounds with the reputation-free names (treatment 1). The other 10 groups played 10 PG rounds with the reputation-free names, followed by 5 rounds with the reputation names (treatment 2). The students were not informed about the number of rounds to be played, the sequence of IR and PG rounds or the sequence of pseudonyms to be used.

3. Results

In the groups of treatment 1 (PG rounds 11 to 20 with the reputation name and PG rounds 21 to 25 with the reputation-free name) the level of cooperation was significantly higher during the rounds with the reputation name (average cooperation per round 63.0%) than during the rounds with the reputation-free name (average cooperation per round 43.7%), (Wilcoxon signed ranks matched pairs test, z=1,99, p=0.047, n=10 groups, two tailed; we use each group of 6 students as the statistical unit to avoid pseudoreplication) (fig. IV.1). In the groups of treatment 2 (PG rounds 11 to 20 with the reputation-free name and PG rounds 21 to 25 with the reputation name) the level of cooperation was also significantly higher in rounds with the reputation name (average cooperation per round 66.0%), than in rounds with the reputation-free name (average cooperation per round 48.2%, Wilcoxon signed ranks matched pairs test, z=2.60, p=0.009, n=10 groups, two tailed). Combining the probabilities from treatment 1 and 2 depicts a highly significant overall effect (Fischer combination test, chi-square=15.538, p<0.005, df=4).

The players of treatment 1 earned significantly more money (average 1.74 € per round) with the reputation name in PG rounds than with the reputation-free name (average 1.06 €) (Wilcoxon signed ranks matched pairs test, z=2.80, p=0.005, n=10, two tailed). This was also the case in treatment 2 (average 1.65 € per round with reputation name, 1.20 € per round with reputation-free name, Wilcoxon signed ranks matched pairs test, z=2.60, p=0.009, n=10, two tailed). The combined effect is highly significant (Fisher combination test, chi-square=20.018, p<0.001, df=4).

4. Discussion

This study shows that the knowledge of being recognized with the same identity (name) in both a public goods (PG) game and an indirect reciprocity (IR) game produces a high level of cooperation in the public goods game. When the subjects had different identities in the two games, cooperation in the public goods game declined as usual. This result implies that humans are well aware of whether they will be recognized in a future social situation, and use this information to invest in their reputation only if it will probably pay off in the other context. This can be called strategic investment in reputation. Similarly, when human subjects were allowed to punish uncooperative players (by imposing a fine) in a public goods game, the level of cooperation increased immediately (Fehr and Gächter, 2002) showing that they were well aware of whether uncooperative behavior could be punished.

We cannot exclude that the interaction of playing the two games has an effect on the level of cooperation in addition to subjects having always the same identity; such an additional interaction effect is suggested by the relatively high level of cooperation in PG rounds with the reputation-free name.

Our results also imply that being recognized can remove the social dilemma from a public goods scenario. If everybody is under pressure to invest in her reputation, there is no longer a conflict between the group's and the individual's interest. The public resource was almost maximized and we found that everybody gained a high personal payoff in this scenario. The payoff per player was significantly higher when the subjects knew they would be recognized in the other game than when they expected to be unrecognizable there. Reputation may be a currency that can be used in various social contexts (Sigmund et al., 2001). In a previous study (Milinski et al., 2002a, chapter I) we found that donations made to charity (UNICEF) significantly increased the probability of being helped in an indirect reciprocity game if the donations were made public. This shows that people can actively invest in their reputation even when no public goods situation is available.

However, even full anonymity does not necessarily cause the breakdown of cooperation in a public goods situation if certain conditions are met. Recently, Hauert et al. (2002b) proposed that with optional participation in the public goods game “loners”, i.e. those players who do not join a public goods group, cooperators who join the group and contribute into the public good and defectors who join the group but do not contribute into the public good will coexist through a rock-paper-scissors dynamics even under full anonymity. An experimental study (Semmann et al., 2003, chapter III) showed that the opportunity for this kind of „volunteering“ easily generates this dynamics in public goods games with human subjects. The rock paper scissors dynamics comes up, because if the majority of the group chooses to be cooperators the highest payoff is achieved as a defector, resulting in an increasing number of players choosing this strategy. When defectors dominate, the highest payoff can be achieved through choosing to be a loner. Finally when loners dominate the public good group size is very small and the highest payoff is achieved by choosing to cooperate in such small groups. However when cooperators increase in numbers the public goods group size increases again and the cycle continues. Through the recurring rise of loners, cooperators, defectors and the connected public goods group size changes cooperation is perpetuated at a substantial level on average. However, the results of the present study suggest that if anonymity would be removed the decisions made in any public goods situation could be recalled in other social games and would thereby be connected to reputation building, the rock-paper-scissors dynamics is reduced and may eventually disappear. As a result cooperation would be perpetuated at an even higher level even in larger public goods groups. This prediction awaits experimental testing.

Acknowledgements

We thank the students at the universities of Kiel, Bonn and Hamburg for their participation; T. Bakker, H. Brendelberger, E. Heinz and K.-P. Sauer for support and the Max-Planck-Institute of Meteorology Hamburg for hospitality.


Chapter V

V REPUTATION IS VALUABLE INSIDE AND OUTSIDE THE OWN SOCIAL GROUP

Abstract

Humans cooperate successfully in groups of unrelated individuals and by doing so pose one of the major problems for the behavioral sciences. The quest for finding the underlying mechanisms has centered on social dilemmas, such as the public goods game, where humans often fail to sustain a public resource. The need to maintain a good reputation for other social interactions has been identified as an effective mechanism to lead to cooperative behavior in public goods situations. Here we show, that building a good reputation in a public goods situation is not only valuable while interacting in other social situations within the own social group. The reputation is also highly valuable when interacting with members of a different social group. Humans reward an individual's good reputation without ever having experienced the individuals' positive behavior themselves. In this experiment humans sustain public resources in order to profit from their good reputation in future encounters with others in and outside the own social group.

Introduction

Modern human societies strongly depend on cooperative behavior, which in many respects is still an unsolved puzzle. The question why unrelated individuals cooperate has entangled scientist since the proposal of reciprocal altruism as one of the underlying mechanisms that promote cooperative behavior (Trivers, 1971). Especially the problem why humans are not able to sustain a public resource, which everybody in a group of genetically unrelated individuals is free to overuse, has been a major focus in the past. Hardin first described this situation as the "tragedy of the commons" (Hardin, 1968) and it has been studied intensively ever since as the public goods game (Berkes et al., 1989; Dawes, 1980; Hardin, 1998; Ledyard, 1995; Ostrom, 1999; Ostrom et al., 1999). Recently empirical and theoretical studies have identified reputation building (Milinski et al., 2002b, chapter II), punishment (Boyd and Richerson, 1992; Fehr and Gächter, 2000; Gintis, 2000b), even altruistic punishment (Fehr and Gächter, 2002) and volunteering (Hauert et al., 2002b; Semmann et al., 2003, chapter III) as mechanisms that promote cooperation in this type of social dilemma. The standard public goods game (Ledyard, 1995) consists of four players who have the possibility to contribute a money unit into a public pool anonymously. The content of the pool is then doubled and paid evenly to all players irrespectively of their contributions into the public pool. This situation poses a conflict between the group interest and the individuals interest. The group stands best if all players contribute into the public pool. However, the rational individual should never invest into the public pool, because a money unit invested is doubled and then divided by four players. Therefore only half of the money unit is returned to original investor. In addition defectors do always better than cooperators within the same group. Nevertheless experiments with humans usually begin with a high cooperative level that declines over time (Fischbacher et al., 2001). In order to avoid the decline of cooperation one can make the decisions in the public goods situation reputation relevant by making these decisions known in an indirect reciprocity context (Milinski et al., 2002b, chapter II). In indirect reciprocity situations a good reputation (Alexander, 1987; Zahavi, 1991) or positive image score (Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b) can be achieved by acts of help and experimental studies have shown that a good image score is rewarded by others (Milinski et al., 2001; Seinen and Schram, 2001; Wedekind and Milinski, 2000). A recent study confirmed that a good reputation gained in public goods situations is also highly valuable in indirect reciprocity games (Semmann et. al. submitted, chapter IV).

Here we present that building a good reputation in public goods games is potentially valuable outside the own social group where the reputation was built. A good reputation is possibly just like a currency transferable to a different social group, where it valued just as high as within the own social group. Humans regularly interact with numerous people in different contexts. With some people one only interacts in one social context or group whereas with others one interacts in more than one social context or group. Nevertheless it is not unlikely that information about the interactions in other social groups is accessible. For example picture a person A has two neighbors B and C and all three work for the same corporation. However, neighbor B is in the same department and neighbor C is not. So our person A is interacting with neighbor C exclusively on a private basis and with neighbor B professionally and privately. Nevertheless person A is able to hear about neighbor Cs behavior at the company or might even observe the interactions himself without ever being involved. Therefore person A has the possibility to take the behavior of neighbor C at the company into account when deciding whether or not to help in a private context.

Method

We conducted the experiment with 228 Students of the Universities of Hamburg, Bonn and Kiel. Always 12 Students formed a group, which played a computerized game. The participants made their decisions throughout the game anonymously in regard to their real life identity. Several measures were necessary to preserve the players anonymity while the game was in progress. All participants were separated by partitions and could make their decisions through silent yes and no buttons at their desk. A short oral introduction included an explanation of how the individuals' anonymity was assured and the use of the buttons. This was followed with choosing a connecting cable from a bunch of similar looking cables, which were disconnected after the game to furthermore assure the participants of the anonymity of their decisions. The necessity to carry out the experiment double blind has been previously pointed out by Hoffman et.al. (1996) as a way to prevent an artificial rise of cooperation. This procedure was followed by a detailed text introduction by the computer ending with the program assigning each player a pseudonym. The pseudonyms were names of moons of our solar system (e.g. Telesto, Nereid) in order to have memorable names without possible prefixed reputations. The players could observe the complete game over a large screen on the wall. Every decision made by a player throughout the game was shown together with the players' pseudonym. Therefore the decisions of the players were not anonymous, but associated with their pseudonym.

PG groups	1 2 3 4 5 6	A B C D E F
IR groups	1 2 3 D E F	A B C 4 5 6

Fig. V.1: Group composition in PG and IR rounds.

In each public goods (PG) round the 12 participants were divided in two PG groups of six players (e.g. in the first PG group were the players 1 to 6 and in the second PG group were the players A to F; Fig. V.1). The composition of the two PG groups was maintained throughout the experiment. Each PG round consisted of the program announcing that it was now the first six players (1-6) turn to play. The pseudonyms of these players were shown and they were asked if they wanted to contribute 1.25 € into their public pool. When all players had made their decision the results were displayed on the screen for 25 seconds. The same procedure was then carried out with the second six players (A-F). The composition of the PG groups was always the same in all PG rounds.

For all indirect reciprocity (IR) rounds three players were exchanged between the PG groups (e.g. in the first IR group were the players: 1, 2, 3, D, E, F and in the second IR group were the players: 4, 5, 6, A, B, C; Fig. V.1). The composition of the two IR groups was also maintained for all IR rounds throughout the experiment. Therefore each player (e.g. player 1 in Fig. V.1) had two players he was interacting with in both situations (players 2 and 3), three players he would exclusively meet in PG rounds (players 4, 5 and 6) and three players he would only interact with in IR rounds (players D, E and F). An indirect reciprocity round started with the announcement that the first group of six players would now play the IR round. Throughout the round each player was once in the role of the potential donor and once in the role of the potential receiver. As a donor the player was asked if she wanted to give 1.25 € from her account to the receiver. If the player decided YES the money was taken from her account and 2.00 € were credited to the account of the receiver. The amount credited to the receivers account was higher, because in theory, the donors costs of giving help is lower than the value of the help to the receiver (see (Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b). Before a player made her decision all decisions of the receiver in the past rounds were displayed. Finally the donors' decision was displayed for three seconds. The players knew from the introduction that two players who have met once with certain roles (e.g. player A as donor and player B as receiver) would never meet again throughout the game in alternated roles (e.g. player B as donor and player A as receiver). They knew it was possible that they would meet with the same roles again. After the first six players had been all once donor and once receiver it was announced that the second group of six players would now play the IR round.

The experiment lasted 16 rounds. The first round was an IR round followed by alternating PG rounds and IR rounds until round 12. The last four rounds were exclusively PG rounds. At the end of round 12 the players were informed that only PG rounds would follow until the end of the game. The participants had no knowledge of the number of rounds to be played at any time throughout the experiment. After the experiment the participants were asked to answer the following question on a questionnaire: “Where you aware while playing with someone in a pair round if you had also played with this person in the group round?” (The terms used in the introduction and the questionnaire were pair round instead of IR round and group rounds instead of PG round). If they answered this question with yes they could also clarify this by marking always, often, seldom or never. The participants were furthermore asked to clarify their decision and were free to make additional comments.

Results

The average cooperation was fairly stable on a very high level until round 12, where the players were informed that only public goods rounds would follow (Fig. 2).

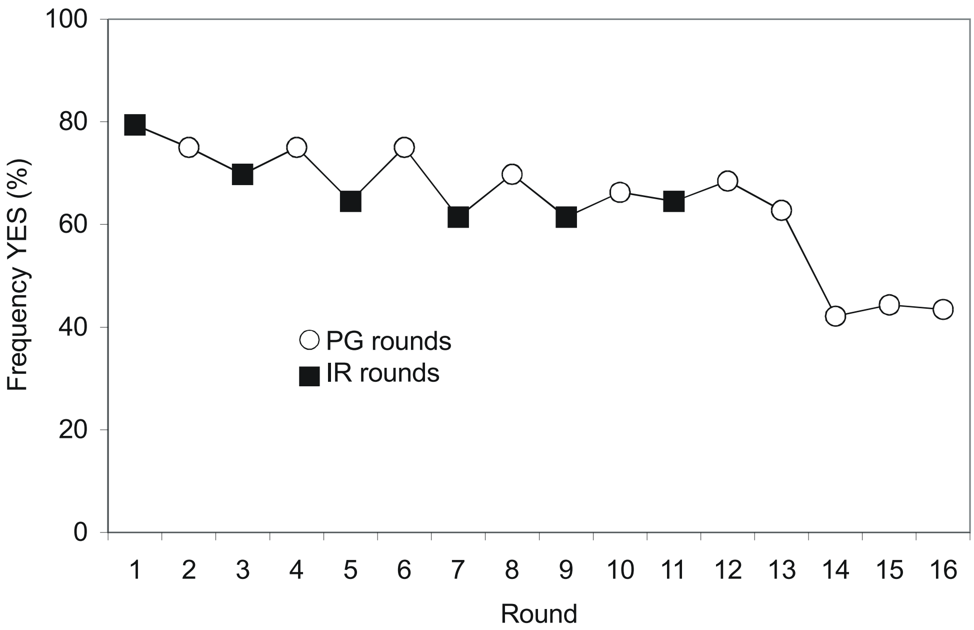


Fig. V.2 Average Cooperation per round per group. Black square symbols represent IR rounds and white round symbols represent PG rounds.

The initial cooperation in round 1 was very high (79.39 %) and dropped significantly from to round 3 (69.73 %)(compared average cooperation round 1 and round 3, Wilcoxon signed rank test, n = 19, p = 0.013, Z = -2.50, two tailed; Fig. 2; to avoid pseudoreplication we used each group (19 groups) of twelve subjects as our statistical unit in the first part of our analysis). Thereafter the cooperation is fairly stable and the decline was not significant any more (compared average cooperation in round 3 and round 11, Wilcoxon signed rank test, n = 19, p = 0.33, Z = -9.74, two tailed). After the announcement that only PG rounds would follow until the end of the game in round 12 the cooperation declined highly significantly (compared average cooperation in round 12 and round 16, Wilcoxon signed rank test, n = 19, p = 0.001, Z = -3.31, two tailed). As in previous experiments the cooperation until round 12 was higher in PG rounds than in indirect reciprocity rounds (compared average cooperation of six PG rounds and six IR rounds 12 and round 16, Wilcoxon signed rank test, n = 19, p = 0.008, Z = -2.66, two tailed).

Players who had refused to give in a PG round where more likely to receive NO in the following IR round than players who had given in a PG round (paired t-test, n = 19, p <= 0.0001, T = -5.93, two tailed). If the players were playing in the same PG group the probability to receive NO in the next IR round after refusing to give in the PG round was still significantly higher than after giving in the PG round (paired t-test, n = 19, p = 0.002, T = -3.532, two tailed; Fig. 3). However, to have played together with the donor in the same PG group did not further increase the likelihood to receive NO compared to having a donor who played in the other PG group (paired t-test, n = 19, p = 0.743, T = -0.333, two tailed; Fig. 3). If the players were in different PG groups the probability to receive NO in the next IR round after refusing to give in the PG round was yet again significantly higher than after giving in the PG round (paired t-test, n = 19, p = 0.01, T = -2.900, two tailed; Fig. 3). And once more not to have played together with the donor in the same PG

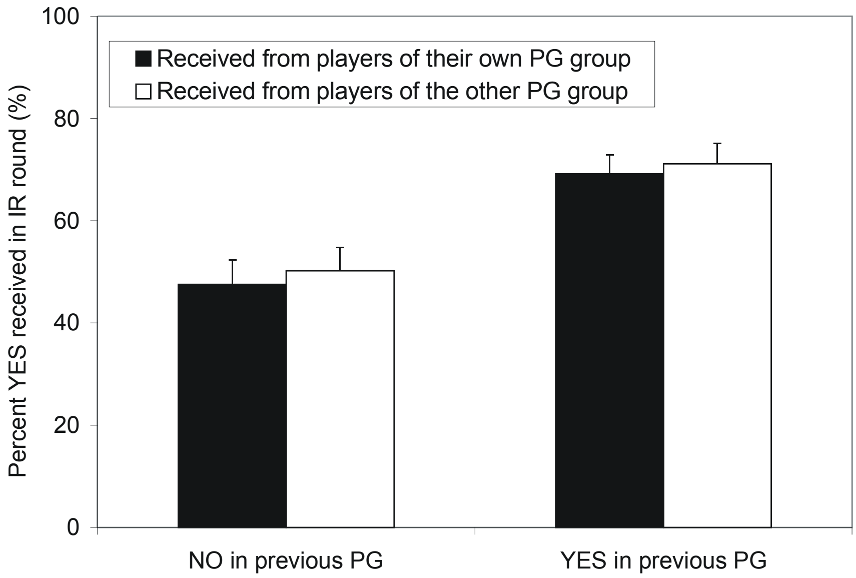


Fig. V.3 Probability to receive YES in a round depending on whether the recipient had either given or not given in the previous PG round and on whether the recipient and the donor had played together or not together in the same PG group. The probability per group (mean ± s.e.m.) for all four situations is shown.

group did not increase the likelihood to receive NO compared to having a donor who played in the same PG group (paired t-test, n = 19, p = 0.542, T = -0.622, two tailed; Fig. 3).

In order to confirm the negative finding we tested whether it really had no influence on the decision to having played with someone in the same public goods group or not. Since there was no difference on the group level to confirm this, the conservative approach was to consider each decision of every player as the statistical unit in order to use all available information to find a difference in the behavior. The analysis was done with a binary logistic regression model in order to use the complete available information even though we increased thereby pseudoreplication to some extend in comparison to using the group as the statistical unit. Nevertheless there was again no difference between having played together in public goods rounds or not and the probability to receive YES in an IR round (binary logistic regression, n = 1140, p=0.820, β= 0.588, s.e.=0.062).

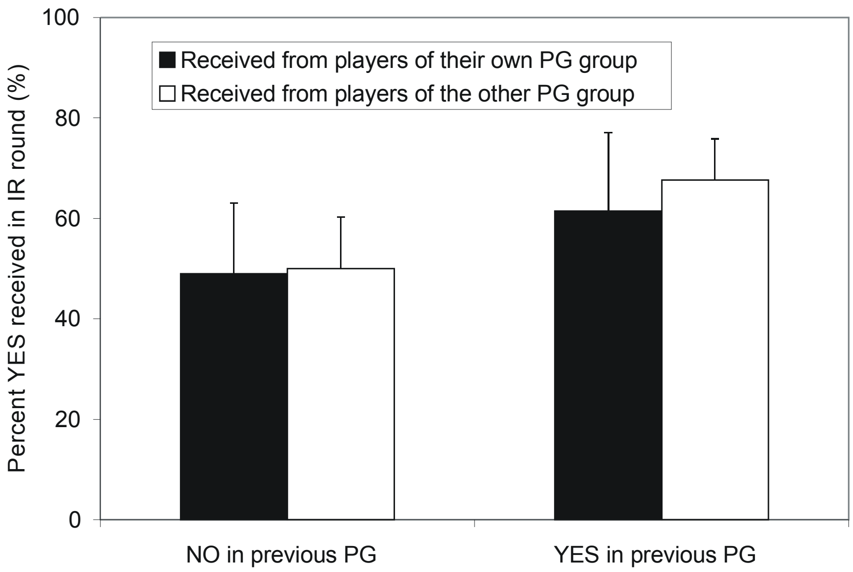


Fig. V.4 Probability to receive YES in a round depending on whether the recipient had either given or not given in the previous PG round and on whether the recipient and the donor had played together or not together in the same PG group. The probability per group (mean ± s.e.m.) for all four situations is shown. Only the decisions are included of participants who had answered in the questionnaire that they were conscious about the fact whether or not they had played together with the current receiver in the PG round and in addition had also said they had used this information for their decision finding always or often.

Consequently going one step further we incorporated the results of the questionnaires into the analysis. The questionnaires showed that of 228 Students 103 (45.18%) answered that they were not aware throughout the game whether they had played with the current pair rounds (IR rounds) partner also in the group rounds (PG round). The other 115 Students (54.82%) answered that they indeed where aware of this fact. Of these 115 Students, 3 specified that they had always used this information, 35 used this information frequently, 58 used this information seldom and 18 did not use this information. The remaining 11 students did not specify their positive answer further. We again analyzed the data, but this time we included only the decisions of those players who had answered the questionnaires with YES in combination with either always or often. If there was any difference between the treatment of players of the own PG group and the other PG group then this should show here. However there was still no difference (binary logistic regression, n=170, p=0.911, s.e.=0.156, β=0.357). Nonetheless the difference of the probability to receive YES in an IR round whether a player had given or not given in the previous PG round remained significant (binary logistic regression, n=170, p=0.017, s.e.=0.156, β=0.357)

Discussion

There is good empirical evidence that building a good reputation is highly valuable for interacting with members of the own social group. A good reputation increases the chance of receiving help from others (Milinski et al., 2002b, chapter II; Seinen and Schram, 2001; Wedekind and Milinski, 2000) and additionally even increases the likelihood to be elected in a political context (Milinski et al., 2002a, chapter I). Furthermore it has been shown that humans strategically invest to preserve their good reputation (Semmann et. al. submitted, chapter IV).

Members of the own social group often times profit directly or indirectly from building a good reputation and reward such behavior with cooperation. This study shows that even outsiders who have never profited of the process of reputation building reward a good reputation gained outside their social group. There are two facts that indicate that the players do not treat group members and outsiders differently. First we found that the average cooperation was higher in public goods rounds than in the indirect reciprocity rounds, resembling the results of our previous study with only one social group (Milinski et al., 2002b, chapter II). Secondly the majority of participants answered in the questionnaire that they either paid no attention to the available information or simply did not use it. Directly testing for whether players behave differently towards group members and outsiders, we found no difference on the group level or the single decision level, even for those players who had explicitly answered in the questionnaire that they used this information, treated both types of players equally.

In order for this mechanism to work the information about the individuals' reputation has to be accessible over the boarders of the social group. This can be achieved through observation of interactions and possibly through gossip or written records. Reputation can be understood as a valuable currency, which can be accumulated in social interactions. A good reputation is gained by playing by the rules of the social community. Similar to other currencies our data shows that it is transferable between social groups. This transferability might be limited to similar cultures, because a good reputation, which was built under different rules for socially accepted behavior, might not transfer as easily. Even though that person has played by the rules of one cultural context it might not resemble acceptable behavior in a different culture. For instance the table etiquette is quite different between cultures, it is not acceptable to eat with your fingers in western cultures whereas it is perfectly acceptable in other cultures.

Reputation building seems to have a major impact on the stability of cooperative behavior in humans. Transferability to other social groups may be one of the essential mechanisms necessary to sustain cooperation in very large social groups in which individuals interact often with some members and very seldom with others.

Acknowledgements

We thank the students at the universities of Kiel, Bonn and Hamburg for their participation; T. Bakker, H. Brendelberger, E. Heinz and K.-P. Sauer for support, the Max-Planck-Institute of Meteorology Hamburg for hospitality and Inke König for help with the statistical modeling.

Conclusion

Conclusion

When I started working on human cooperative behavior it seemed obvious to me why humans cooperate. My view of cooperation was that everybody profits when people work together for a common goal. This simple view was most likely the result of growing up in a society, which is strongly influenced by cooperation between unrelated individuals. Thereby I had learned the rules how and when to cooperate with other by heart. However, the ever present, oftentimes subtle conflict between the individuals’ interest and the group interest discussed in the previous five chapters reveals that there is much more involved in cooperative behavior. This work altered my original view of cooperation and sharpened my eye for conflict situations.

Each chapter shows us a piece of the puzzle of why and under which circumstances cooperative behavior can be established and sustained. Introducing reputation can change the outcome of usually uncooperative situations. It is a very strong mechanism that leads to cooperative behavior in humans (chapter II), by aligning the group interest with the individuals’ interest. In order to achieve a good reputation one can help others in the own social group. However, it is also possible to increase ones reputation through giving help to charity, even though the help is thereby given to others outside the own social group (chapter I). An individuals’ good reputation can be viewed as a currency that can be accumulated by helping others. Humans do not differentiate where the reputation was built; rather this currency is highly valuable within and outside the own social group (chapter V).

Nevertheless achieving a good reputation is costly. If reputation is missing, as the incentive, humans reduce their investment into public resources. Humans are very aware of the situational circumstances and very strategically invest into their reputation (chapter IV). However, even in complete absence of reputation or other mechanisms that promote cooperation, like punishment, it is possible to establish circumstances under which cooperation can be sustained, on average on a substantial level (chapter III).

There are still many open questions about human cooperative behavior. Not all acts of help have the same value, so what kind of value system do we apply? How exactly do we keep track of the reputation of social partners? Do we keep an accumulated score or do we assess a score in the beginning of a relationship and later on only adjust the score when the partners' behavior changes? Future studies should also discriminate between individuals. Humans often behave quite different in exactly the same circumstances, probably depending on previous experience, knowledge and at least to some extend genetics. In our experiments we could observe a wide range of individual behavior, ranging from highly cooperative individuals to unconditional defectors. Therefore also the mixture of groups can be viewed as a factor that influences whether cooperation can be achieved or not. It will be crucial to know in the future these and other unknown circumstances, in order to predict how circumstances have to be altered, to make people behave more cooperatively.

Danksagung

Allen voran möchte ich Manfred Milinski, für die exzellente Betreuung und die Möglichkeit in seiner Abteilung am Max-Planck-Institut Plön meine Doktorarbeit durchzuführen, danken.

Ebenfalls möchte ich den folgenden Personen danken, ohne deren Mithilfe diese Arbeit nicht möglich gewesen wäre: Hans-Jürgen Krambeck für die immerwährende Unterstützung bei der Programmierung der Versuchsanlage und allen anderen computerrelevanten Problemen, Herbert Kiesewetter für die technische Unterstützung und den Zusammenbau der computergesteuerten Hardware, Derk Wachsmuth für das gemeinsame Erlernen der Computersprache Java und die Hilfe in bezug auf PC Probleme, Werner Wegner für technische – und Software – Unterstützung, Harald Deiwick für seine Hilfe beim Bau unserer Versuchsanlage und August Hämmerli für das Korrekturlesen und die Layoutvorlage als Leitfaden für diese Arbeit.

Des weiteren danke ich den Helfern an den Universitäten Hamburg, Kiel und Bonn. Theo C. M. Bakker gilt mein Dank für die Unterstützung bei der Rekrutierung der Studenten für unsere Experimente und Hilfe vor Ort. Dem Max-Planck-Institut für Meteorologie Hamburg danke ich, für die Möglichkeit ihren Seminarraum für unsere Experimente zu nutzen.

Allen meinen Freunden am Max-Planck-Institut möchte ich für die Hilfe während meiner Zeit am MPI Plön und den körperlichen und seelischen Ausgleich (besonders zu stressigen Zeiten) danken. Skifahren auf dem Bungsberg wird wohl immer als Highlight in meiner Erinnerung bleiben...

Meiner Familie danke ich für ihre uneingeschränkte Unterstützung. Zu guter Letzt danke ich meiner Frau Karen für den persönlichen Beistand und die praktischen kleinen Hilfen.

References

Alexander R. D. 1987 The Biology of Moral Systems. New York, de Gruyter.

Axelrod R. 1984 The evolution of cooperation. New York, Basic Books.

Axelrod R., Hamilton W. D. 1981 The evolution of cooperation. Science 211, 1390-1396.

Berkes F., Feeny D., McCay B. J., Acheson J. M. 1989 The benefits of the commons. Nature 340, 91-93.

Bolton G. E., Katok E., Ockenfels A. 2001 What´s in a reputation? Indirect reciprocity in an image scoring game. Working paper, University of Wisconsin.

Boyd R., Richerson P. J. 1988 The evolution of reciprocity in sizable groups. Journal of Theoretical Biology 132, 337-356.

Boyd R., Richerson P. J. 1992 Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology 13, 171-195.

Broadbent D. E. 1965 Information processing in the nervoussystem. Science 150, 457-462.

Brown J. L. 1983 Cooperation - a biologist´s dilemma. Advances in the Study of Behavior 13, 1-37.

Christen M., Milinski M. 2003 The consequences of self-fertilization and outcrossing of the cestode Schistocephalus solidus in its second intermediate host. Parasitology 126, 369-378.

Coleman J. S. 1990 Foundations of Social Theory. Cambridge, MA, Havard Univ. Press.

Colman A. M. 1995 Game Theory and Its Applications in the Social and Biological Sciences. Oxford, UK, Butterworth-Heinemann.

Conan Doyle A. 1986 The man with the twisted lip. In: Sherlock Holmes - the complete illustrated short stories. London, Chancellor Press; 94-112.

Connor R. C. 1995 The benefits of mutualism: a conceptual framework. Biological Review 70, 427-467.

Dawes R. M. 1980 Social dilemmas. Annual Review of Psychology 31, 169-193.

Fehr E., Gächter S. 2000 Cooperation and punishment in public goods experiments. American Economic Review 90, 980-994.

Fehr E., Gächter S. 2002 Altruistic punishment in humans. Nature 415, 137-140.

Fehr E., Rockenbach B. 2003 Detrimental effects of sanctions on human altruism. Nature 422, 137-140.

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

Gintis H. 2000a Game Theory Evolving. Princeton, NJ, Princeton Univ. Press.

Gintis H. 2000b Strong reciprocity and human sociality. Journal of Theoretical Biology 206, 169-179.

Gintis H., Smith E. A., Bowles S. 2001 Costly signaling and cooperation. Journal of Theoretical Biology 213, 103-119.

Hamilton W. D. 1964a Genetical evolution of social behaviour. Journal of Theoretical Biology 7, 1-52.

Hamilton W. D. 1964b Genetical Evolution of Social Behaviour 2. Journal of Theoretical Biology 7, 17-&.

Hardin G. 1968 Tragedy of commons. Science 162, 1243-1248.

Hardin G. 1998 Extensions of "the tragedy of the commons". Science 280, 682-683.

Hauert C., De Monte S., Hofbauer J., Sigmund K. 2002a Replicator dynamics for optional public good games. Journal of Theoretical Biology 218, 187-194.

Hauert C., De Monte S., Hofbauer J., Sigmund K. 2002b Volunteering as red queen mechanism for cooperation in public goods games. Science 296, 1129-1132.

Hoffman E., McCabe K., Smith V. L. 1996 Social distance and other-regarding behavior in dictator games. American Economic Review 86, 653-660.

Ledyard J. O. 1995 In: Handbook of Experimental Economics (Kagel JH, Roth AE, eds). Princeton, NJ, Princeton Univ. Press; 111-194.

Leimar O., Hammerstein P. 2001 Evolution of cooperation through indirect reciprocity. Proceedings of the Royal Society of London Series B-Biological Sciences 268, 745-753.

Lotem A., Fishman M. A., Stone L. 1999 Evolution of cooperation between individuals. Nature 400, 226-227.

Magurran A. E., Pitcher T. J. 1987 Provenance, shoal size and the sociobiology of predator-evasion behavior in minnow shoals. Proceedings of the Royal Society of London Series B-Biological Sciences 229, 439-465.

Michor F., Nowak M. A. 2002 Evolution - the good, the bad and the lonely. Nature 419, 677-679.

Milinski M. 1990 Information overload and food selection. In: Behavioral mechanisms of food selection (Hughes RN, ed). Berlin, Springer-Verlag; 721-737.

Milinski M., Semmann D., Bakker T. C. M., Krambeck H. J. 2001 Cooperation through indirect reciprocity: image scoring or standing strategy? Proceedings of the Royal Society of London Series B-Biological Sciences 268, 2495-2501.

Milinski M., Semmann D., Krambeck H. J. 2002a Donors to charity gain in both indirect reciprocity and political reputation. Proceedings of the Royal Society of London Series B-Biological Sciences 269, 881-883.

Milinski M., Semmann D., Krambeck H. J. 2002b Reputation helps solve the 'tragedy of the commons'. Nature 415, 424-426.

Milinski M., Wedekind C. 1998 Working memory constrains human cooperation in the Prisoner's Dilemma. Proceedings of the National Academy of Sciences of the United States of America 95, 13755-13758.

Mohtashemi M., Mui L. 2003 Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism. Journal of Theoretical Biology 223, 523-531.

Nowak M. A., Sigmund K. 1998a The dynamics of indirect reciprocity. Journal of Theoretical Biology 194, 561-574.

Nowak M. A., Sigmund K. 1998b Evolution of indirect reciprocity by image scoring. Nature 393, 573-577.

Nowak M. A., Sigmund K. 2000 Social science - Shrewd investments. Science 288, 819-820.

Ostrom E. 1999 Governing the Commons. Cambridge, Cambridge Univ. Press.

Ostrom E., Burger J., Field C. B., Norgaard R. B., Policansky D. 1999 Sustainability - Revisiting the commons: Local lessons, global challenges. Science 284, 278-282.

Parker G. A., Chubb J. C., Roberts G. N., Michaud M., Milinski M. 2003 Optimal growth strategies of larval helminths in their intermediate hosts (vol 16, pg 47, 2003). Journal of Evolutionary Biology 16, 362-362.

Rainey P. B., Rainey K. 2003 Evolution of cooperation and conflict in experimental bacterial populations. Nature 425, 72-74.

Ridley M. 1996 The Origins of Virtue. London, Penguin Books.

Roberts G. 1998 Competitive altruism: from reciprocity to the handicap principle. Proceedings of the Royal Society of London Series B-Biological Sciences 265, 427-431.

Schelling T. C. 1973 Hockey helmets, concealed weapons, and daylight saving - study of binary choices with externalities. Journal of Conflict Resolution 17, 381-428.

Seinen I., Schram A. 2001 Social status and group norms: indirect reciprocity in a helping experiment. Discussion Paper TI2001-003/1. Amsterdam, Tinbergen Institute.

Selten R., Ockenfels A. 1998 An experimental solidarity game. Journal of Economic Behavior & Organization 34, 517-539.

Semmann D., Krambeck H. J., Milinski M. 2003 Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature 425, 390-393.

Sigmund K., Hauert C., Nowak M. A. 2001 Reward and punishment. Proceedings of the National Academy of Sciences of the United States of America 98, 10757-10762.

Sober E., Wilson D. S. 1999 Unto Other: The Evolution and Psychology of Unselfish Behavior. Cambridge, MA, Havard Univ. Press.

Sugden R. 1986 The Economics of Rights, Co-operation and Welfare. Oxford and New York, Blackwell.

Trivers R. 1985 Social Evolution. Menlo Park, CA, Benjamin Cummings.

Trivers R. L. 1971 Evolution of Reciprocal Altruism. Quarterly Review of Biology 46, 35.

Velicer G. J., Yu Y. T. N. 2003 Evolution of novel cooperative swarming in the bacterium Myxococcus xanthus. Nature 425, 75-78.

Wedekind C., Milinski M. 2000 Cooperation through image scoring in humans. Science 288, 850-852.

Wilson D. S., Sober E. 1994 Reintroducing Group Selection to the Human Behavioral-Sciences. Behavioral and Brain Sciences 17, 585-608.

Zahavi A. 1991 In: Cooperative Breeding in Birds: Long Term Studies in Behaviour and Ecology (Stacey PB, Koenig WD, eds). Cambridge, Cambridge Univ. Press; 105-130.

Zahavi A. 1995 Altruism as a handicap - the limitations of kin selection and reciprocity. Journal of Avian Biology 26, 1-3.

Glossary

Direct fitness (Direkte Fitneß)

An individuals’ genetic contribution to the next generation through the own reproduction.

Evolutionary stable strategy (Evolutionsstabile Strategie, ESS)

Strategies that cannot be replaced by alternative strategies, when a certain proportion of individuals within a population uses these strategies.

Fitness (Fitneß)

An individuals’ overall genetic contribution to the next generation. Often measured in the number of surviving grandchildren.

Group selection (Gruppenselektion)

The process when groups differ in their collective characteristics and this difference correlates with the survival chances of the group.

Kin selection (Verwandtenselektion)

The process when individuals differ in a way that influences parental care or their helping behavior, thereby changing the survival chances of their offspring or other relatives.

Mutualism (Mutualismus)

A relationship or an action that is profitable for both parties involved.

Reciprocal altruism / Reciprocity (Reziproker Altruismus / Reziprozität)

Help is given by a donor, which the recipient of the help returns at a later point in time.

Appendix

Original text introduction by the computer program (experiment chapter IV)

Page 1

Willkommen zu diesem Experiment, in dem Sie Geld verdienen können.

Zu Beginn des Experiments bekommen Sie 10,- Euro auf Ihr Konto gutgeschrieben.

Während des Experiments können sie Geld dazugewinnen oder verlieren.

Das hängt von Ihren eigenen Entscheidungen und den Entscheidungen der anderen

TeilnehmerInnen ab. Ihre Entscheidungen sind anonym.

Am Ende wird Ihnen Ihr Kontostand bar ausgezahlt.

Vor Ihnen befindet sich ein Schaltkasten mit zwei Tasten und einer roten Leuchte. Es ist nur möglich über die Tasten eine Eingabe zu machen, wenn die Leuchte an Ihrem Platz aufleuchtet. Es gibt eine JA- und eine NEIN-Taste.

Sie werden im Verlauf dieses Experiments nur JA oder NEIN Entscheidungen treffen.

DAMIT DIESES EXPERIMENT GELINGT, DÜRFEN SIE AUF KEINEN FALL MIT ANDEREN REDEN ODER SICH AUF EINE ANDERE WEISE BEMERKBAR MACHEN!

Wenn Sie diesen Text vollständig gelesen haben, bestätigen Sie dies bitte mit einem Druck auf die JA-Taste, sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.

Wenn alle TeilnehmerInnen mit JA bestätigt haben, erscheint die nächste Seite.

Page 2

Im Verlauf dieses Experiments werden sie Entscheidungen in zwei unterschiedlichen

Situationen treffen.

IN DER ERSTEN SITUATION (=Paarrunde) treffen zwei Teilnehmer in einer Runde

aufeinander.

Hier gibt es zwei verschiedene Rollen, die Ihnen der Computer während

des Experiments wiederholt, aber gleichhäufig zulost.

Einmal sind Sie in der Rolle der potentiellen GeberIn, das andere Mal sind Sie in der

Rolle der potentiellen EmpfängerIn.

Nur die GeberIn hat eine aktive Rolle, in der sie der EmpfängerIn gibt oder nicht gibt.

Es ist ausgeschlossen, daß eine GeberIn und eine EmpfängerIn in umgekehrten

(reziproken) Rollen im Verlauf des Experiments aufeinandertreffen können.

Z.B. Wenn A als potentielle GeberIn auf B als potentielle EmpfängerIn getroffen ist,

wird B nie als potentielle GeberIn auf A als potentielle EmpfängerIn treffen.

Es werden Ihnen für den Verlauf des Experiments pro Person zwei Pseudonamen

zugeteilt.

Sie spielen also unter zwei verschiedenen Namen.

Die Pseudonamen sind Namen von Monden unseres Sonnensystems (Leda, Triton,

Portia, Sinope, Metis, Ananke, Kallisto, Telesto, Japetus, Despina, Galatea, Okeanos,

Elara, Vestia, Rhea und Nereid)

Wenn Sie diesen Text vollständig gelesen haben bestätigen Sie dies bitte mit

einem Druck auf die JA-Taste sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.

Wenn alle Teilnehmer mit JA bestätigt haben erscheint die nächste Seite.

Page 3

FORTSETZUNG: ERSTE SITUATION (Paarrunde)

Als GeberIn wird Ihnen eine Frage gestellt. "Wollen Sie dieser TeilnehmerIn geben?"

Sie können hierauf mit JA oder NEIN antworten.

Entscheiden Sie sich für JA, so werden Ihrem Konto 1,25 Euro abgezogen und der EmpfängerIn werden 2,00 Euro gutgeschrieben. (0,75 Euro werden von uns dazugegeben.)

Entscheiden Sie sich für NEIN, so wird Ihnen nichts abgezogen und der EmpfängerIn nichts gutgeschrieben.

Entscheidung	GeberIn (Gewinn-Verlust)	EmpfängerIn (Gewinn-Verlust)
JA	-1,25 Euro	+2,00 Euro
NEIN	0,00 Euro	0,00 Euro

Wenn Sie diesen Text vollständig gelesen haben, bestätigen Sie dies bitte mit einem Druck auf die JA-Taste, sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.

Wenn alle TeilnehmerInnen mit JA bestätigt haben, erscheint die nächste Seite.

Page 4

FORTSETZUNG: ERSTE SITUATION (Paarrunde)

Bevor Sie Ihre Entscheidung treffen, erscheinen einige Informationen.

Hier können Sie ablesen, wie die jeweilige EmpfängerIn in den vorausgegangenen Runden entschieden hat, als sie sich in der Rolle der GeberIn befand.

Beispiel:

EMPFÄNGERIN "Triton"

Runde	Hat "Triton" gegeben?
1	JA
2	NEIN
3	NEIN

Falls die TeilnehmerIn noch nicht in der Rolle der GeberIn war erscheint folgendes:

EMPFÄNGERIN "Triton"

Runde	Hat "Triton" gegeben?
1	

Wenn Sie diesen Text vollständig gelesen haben bestätigen Sie dies bitte mit einem Druck auf die JA-Taste sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.

Wenn alle TeilnehmerInnen mit JA bestätigt haben erscheint die nächste Seite.

Page 5

IN DER ZWEITEN SITUATION (Gemeinschaftsrunde) treffen alle Teilnehmer der Gruppe gleichzeitig eine Entscheidung.

Sie werden alle gefragt, ob Sie 1,25 Euro in den Gemeinschaftstopf investieren wollen. Alle Spieler entscheiden nun nacheinander. Die Entscheidungen werden jedoch NICHT an die Wand projeziert. Erst wenn alle Spieler entschieden haben, werden die Entscheidungen angezeigt. Danach wird der Betrag im Gemeinschaftstopf verdoppelt und zu gleichen Teilen an alle Spieler ausgezahlt, unabhängig davon ob Sie in den Gemeinschaftstopf eingezahlt haben oder nicht.

BEISPIEL (5 Spieler):

Nereid	Triton	Galatea	Sinope	Phobos
Entscheidung				
JA	NEIN	NEIN	JA	JA
Abzug vom Konto des Spielers				
-1,25 Euro	-0,00 Euro	-0,00 Euro	-1,25 Euro	-1,25 Euro
Auszahlung an alle Spieler (Verdoppelter Gemeinschaftstopf durch Anzahl der Spieler)				
+1,50 Euro	+1,50 Euro	+1,50 Euro	+1,50 Euro	+1,50 Euro

Wenn Sie diesen Text vollständig gelesen haben bestätigen Sie dies bitte mit einem Druck auf die JA-Taste sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.

Wenn alle TeilnehmerInnen mit JA bestätigt haben erscheint die nächste Seite.

Page 6

Es werden Ihnen jetzt ZWEI Pseudonamen zugeteilt.

Ihr erster Pseudoname WIRD IMMER in den Paarrunden angezeigt.
In den Gemeinschaftsrunden wird ENTWEDER Ihr erster ODER Ihr zweiter Pseudoname angezeigt.
Die Informationen in der Paarrunde zeigen nur wie der aktuelle Empfänger unter seinem ersten Pseudonamen bis dahin gespielt hat.

Die Pseudonamen werden gleich nacheinander auf dem Bildschirm erscheinen.
Wenn die Lampe an Ihrem Platz leuchtet, ist der angezeigte Name Ihr erster Pseudoname für dieses Experiment.
Bestätigen Sie diesen Namen bitte indem Sie nacheinander
einmal JA und einmal NEIN drücken.

Wenn die Lampe an Ihrem Platz erneut leuchtet, ist der angezeigte Name Ihr

zweiter Pseudoname für dieses Experiment.
Bestätigen Sie diesen Namen bitte genauso wie den ersten, indem Sie nacheinander einmal JA und einmal NEIN drücken.

Wenn Sie diesen Text vollständig gelesen haben bestätigen Sie dies bitte mit einem Druck auf die JA-Taste sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.
Wenn alle TeilnehmerInnen mit JA bestätigt haben erscheint die nächste Seite.

Page 7

JETZT BEGINNT DAS EXPERIMENT!

Sie werden entweder gefragt, ob Sie in den Gemeinschaftstopf einzahlen wollen (die Pseudonamen aller Spieler werden angezeigt), oder Sie werden gefragt ob Sie einer angezeigten EmpfängerIn 1,25 Euro von Ihrem Konto geben wollen oder nicht (nur der Pseudoname eines Spielers wird angezeigt).

Sie haben ein Startguthaben von 10,- Euro auf Ihrem Konto.

Wenn Sie diesen Text vollständig gelesen haben bestätigen Sie dies bitte mit einem Druck auf die JA-Taste sobald die rote Leuchte an Ihrem Schaltplatz aufleuchtet.
Wenn alle TeilnehmerInnen mit JA bestätigt haben erscheint die nächste Seite.

Curriculum vitae

Name	Dirk Semmann
Geburtstag	01.07.1970
Geburtsort	Hamburg
Staatsangehörigkeit	Deutsch
Familienstand	Verheiratet (seit 25.05.02)

Ausbildung

1988	Mount Desert Island High School (Maine, USA) Honorary Diploma
1988 – 1991	Gymnasium Atheneum (Stade, Niedersachsen) Abitur
1991 – 1992	Wehrpflichtabsolvierung Bundeswehr Ausbildung zum Sanitäter, Krankenwagen- und LKW-Fahrer
1990 – 2001	Basketball Halbprofi beim VfL Stade in der 1. Regionalliga Nord
1992 – 2001	Studium an der Universität Hamburg in der Fachrichtung Diplom Biologie Diplomarbeit am Max-Planck-Institut Plön. Thema: „Versuche zur Kooperation durch Indirekte Reziprozität“ Diplomprüfungen Hauptfach: Zoologie Nebenfächer: Hydrobiologie und Humanbiologie
1991 – 2000	Freiberufliche Arbeit als Grafikdesigner
1994 – 1996	Freiberufliche Arbeit als Webdesigner.
1994 – 2000	Freiberuflicher Vertreter für Textilwaren bei der Firma Fischer Textil.
Seit April 2001	Promotion am Max-Planck-Institut Plön

Veröffentlichungen

Aus der Diplomarbeit:

Milinski M., Semmann D., Bakker T. C. M., Krambeck H. J. 2001 Cooperation through indirect reciprocity: image scoring or standing strategy? Proceedings of the Royal Society of London Series B 268, 2495-2501.

Seit beginn der Doktorarbeit:

Milinski M., Semmann D., Krambeck H. J. 2002a Donors to charity gain in both indirect reciprocity and political reputation. Proceedings of the Royal Society of London Series B 269, 881-883.

Milinski M., Semmann D., Krambeck H. J. 2002b Reputation helps solve the 'tragedy of the commons'. Nature 415, 424-426.

Semmann D., Krambeck H. J., Milinski M. 2003 Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature 425, 390-393.

Vorträge und Kongressteilnahmen

- Universität Hamburg, Deutschland 11.12.00 Vortrag: Kooperation bei Stichlingen: Das Ultimatumspiel
- Workshop „Ausgewählte Themen in evolutionärer Ökologie“ Universität Kiel, Deutschland, Biologische Station Westerhever Leuchtturm 08.-09-06.01. Vortrag: Human cooperation: common goods and indirect reciprocity
- Universität Hamburg, Deutschland, 09.07.01. Vortrag: Menschliche Kooperation durch indirekte Reziprozität
- XXVII Internationale Ethologische Konferenz, Tübingen, Deutschland, 22.-29.08.01
- Graduiertentreffen Regensburg 2002 Vortrag: Human cooperative behavior
- Universität Essen, 23.04.02, Vortrag: Menschliches Kooperationsverhalten
- IX Internationale Behavior Ecology (ISBE) Tagung, Montreal, Kanada, 07.-12.07.02 Poster: Rock-paper-scissors dynamics in human cooperative behavior

- 4. Göttinger Freilandtage, Göttingen, Deutschland, 09.-12.12.03 Vortrag: Volunteering leads to a rock-paper-scissors dynamics as an escape from human social dilemmas

Erklärung

Hiermit versichere ich, dass diese Abhandlung – abgesehen von der Beratung durch meine akademischen Lehrer – nach Inhalt und Form meine eigene Arbeit ist und dass ich keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Die Arbeit hat bisher weder ganz noch zum Teil an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegen. Teile dieser Arbeit wurden als Manuskripte bei Zeitschriften veröffentlicht oder eingereicht, mit Manfred Milinski, Hans-Jürgen Krambeck und Theo Bakker als Koautoren. Kapitel I wurde in Proceedings of the National Royal Society London Series B veröffentlicht. Kapitel II und III wurden in Nature veröffentlicht. Kapitel IV wurde bei Behavior Ecology and Sociobiology eingereicht.

Plön den, 16. Dezember 2003 Dirk Semmann