Behav Ecol Sociobiol (2004) 56:248–252
DOI 10.1007/s00265-004-0782-9

ORIGINAL ARTICLE

Dirk Semmann · Hans-Jürgen Krambeck · Manfred Milinski

Strategic investment in reputation

Received: 22 August 2003 / Revised: 23 February 2004 / Accepted: 9 March 2004 / Published online: 2 April 2004

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 sustain the public resource. However, cooperation declines immediately when individual identities switch from being recognizable to being unrecognizable between the two interacting games.

Keywords Cooperation · Public goods · Conditional strategy · Anonymity

Introduction

Cooperative behaviour 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 behaviour within populations of selfish individuals is usually explained through either kin selection (Hamilton 1964), mutualism or reciprocal altruism (Trivers 1971; Axelrod and Hamilton 1981). Recently, theorists (Nowak and Sigmund 1998a, 1998b; Lotem et al. 1999; Leimar and Hammerstein 2001; Fishman 2003; Mohtashemi and Mui 2003) 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 of returning 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 (Wedekind and Milinski 2000; Bolton et al. 2001; Milinski et al. 2001; Seinen and Schram 2001; Milinski et al. 2002a; Wedekind and Braithwaite 2002). Furthermore, client reef fish approach cleaner fish that they observe cooperating with their current client and avoid cleaners that they observe cheating. As cooperative behaviour towards current clients increases the probability of access to future clients, there is a component of indirect reciprocity (and thus reputation building) present in cleaner-client interactions (Bshary 2002).

In the efforts of individuals to achieve higher benefits and to do as well as they can, the individual attributes that are particularly important in explaining behaviour in social dilemmas include the expectations that individuals have about others' behaviour (trust) and the identities that individuals create which reflect their intentions and norms (reputation) (Ostrom 2003). In that sense, all players carry some sort of reputation reflecting their strategic character (Brandt et al. 2003). Reputations may be transmitted via third parties (gossip) or inferred from direct observation of previous interactions. Reputations can be cultivated when behaviour affects both present and future incremental fitness (Pollock and Dugatkin 1992). This means that investment can be made in trustworthy reputation (Ostrom 2003), if the costly investment has a high probability of paying off in future interactions. However, individuals should stop investing in costly reputation as

Communicated by M. Borgerhoff-Mulder

D. Semmann (✉) · H.-J. Krambeck · M. Milinski
Department of Evolutionary Ecology,
Max-Planck-Institute of Limnology,
24306 Plön, Germany
e-mail: semmann@mpil-ploen.mpg.de
Tel.: +49-4522-763252
Fax: +49-4522-763310

soon as they find out that a future pay-off is unlikely to occur. Thus, strategic investment in reputation is expected to be conditional on whether, e.g., reputation is transmitted to a social situation where it might or might not pay off. Humans meet their neighbours repeatedly in various social interactions and should expect that their reputation would be transferred among interactions. However, when visiting other neighbourhoods or other villages, it may be more rewarding to act uncooperatively in a social dilemma unless gossip finds the way home.

Reputation is, however, usually unimportant in "public goods situations", which are typical social dilemmas where initial cooperation declines after a few rounds (Fischbacher et al. 2001). Social scientists, economists and evolutionary theorists have studied public goods situations extensively (Dawes 1980; Berkes et al. 1989; Ledyard 1995; Hardin 1998; 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 took as much energy from its host as if it were alone (Brown et al. 2002; 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; Gintis 2000; Sigmund et al. 2001; Fehr and Gächter 2002), costly signalling with altruistic acts (Gintis et al. 2001), voluntary participation in the public goods game (Hauert et al. 2002; Semmann et al. 2003) and the interaction with indirect reciprocity situations can help solve the "tragedy of the commons" (Milinski et al. 2002b). In another study, Wedekind and Braithwaite (2002) suggested that costly investment in reputation pays off in a subsequent direct reciprocity game (two-persons prisoner's dilemma), although their result could also be caused by cooperative persons being cooperative in both situations.

In our previous study (Milinski et al. 2002b), 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. A bad reputation from not contributing to the public pool was recognized in the indirect reciprocity game where players refused to support such individuals. However, they supported individuals who had contributed to the public pool. Through this transfer of reputation between 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 that was observed thereafter in these groups suggests that the decisions made in the public goods games were regarded as being no longer relevant for the player's reputation in indirect reciprocity games. This could mean that the subjects had strategically invested in their reputation only when a future pay-off was likely. An alternative explanation for the observed breakdown of cooperation may be 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; during unfair interactions, humans have stronger emotional reactions in the brain (Sanfey et al. 2003), which probably require more channel capacity. 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 some rounds but blocked this transfer in other rounds. 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 different pseudonyms. Each participant received two names of moons of our solar system, e.g. Telesto, Kallisto. One name was used only in public goods rounds, whereas the other name was used in rounds of both games. We could have rendered the players completely anonymous to prevent any transfer of reputation. However, we wanted the players to keep a recognizable personal identity within the game, with interruption of the reputation transfer to be the only treatment effect. 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 the knowledge of being recognized in the other game motivates players to invest in their reputation and thus, as a side-effect, sustain the public resource.

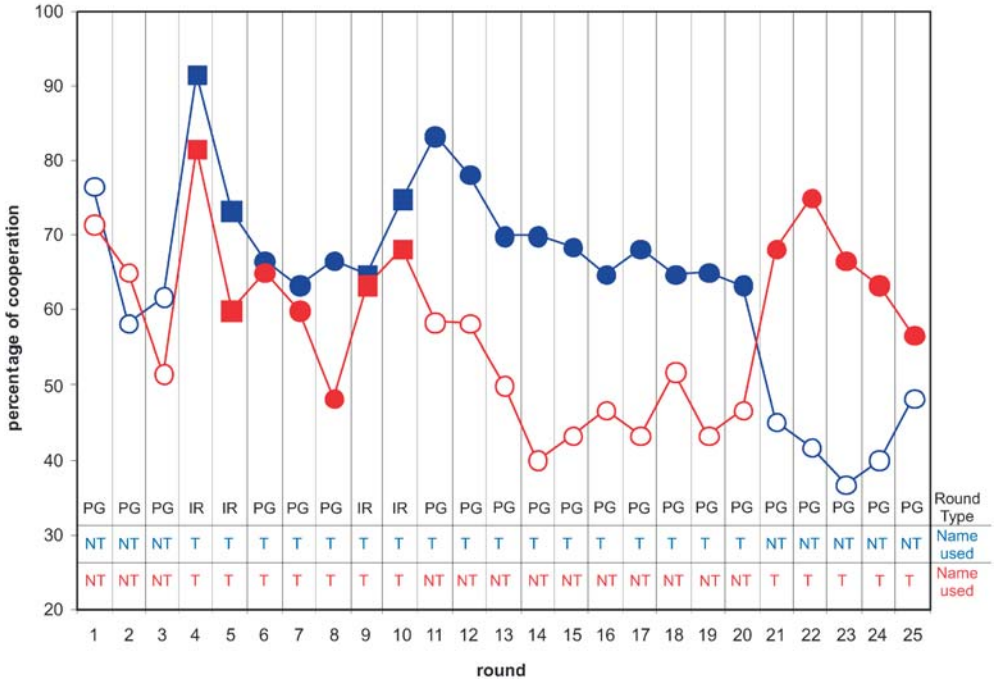

Fig. 1 For the public goods (*PG*) rounds (*circle symbols*) and indirect reciprocity (*IR*) rounds (*square symbols*), the group mean yes per round for both treatments are shown. In treatment 1 (*blue*) the groups played PG rounds, from round 11 to round 20 with their transferable name (*T*) (*filled symbols*) and from round 21 to 25 with their non-transferable name (*NT*). In treatment 2 (*red*) the groups played PG rounds, from round 11 to round 20 with their non-transferable name and from round 21 to 25 with their transferable name. The period from round 1 to 10 was in both treatments identical (three PG rounds played with the non-transferable name, two IR rounds with the transferable name, three public goods rounds with the transferable name and two IR rounds with the transferable name)

Methods

We conducted our experiment with 120 students of the Universities of Bonn, Hamburg and Kiel. Each group consisted of six students, who were anonymous with respect to their real identity but who 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 (see Milinski et al. 2001 for details), 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 (see Hoffman et al. 1996).

The students played a mixture of public goods (PG) rounds, during which all six 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 six 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 was taken from the player's account and €2.00 was 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 made her decision, some information about the potential receiver's behaviour as a donor in earlier rounds was displayed.

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

Each group started with three PG rounds using the non-transferable name, followed by two IR rounds, where the decisions of the first three rounds were not displayed (see Fig. 1). Rounds 6–8 were PG rounds with the transferable name. Rounds 9 and 10 were IR rounds, where all previous decisions made with the transferable name were displayed. This introductory part was the same for both treatments to help the students to become accustomed to the procedure. The last fifteen rounds (11–25) were the actual test, in which the 2 treatments differed in order to control for sequence

effects. Ten groups played ten PG rounds with the players' transferable names, followed by five PG rounds with the non-transferable names (treatment 1). The other ten groups played ten PG rounds with the non-transferable names, followed by five rounds with the transferable 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.

Results

In the groups of treatment 1 (PG rounds 11–20 with the transferable name and PG rounds 21–25 with the non-transferable name), the level of cooperation was significantly higher during the rounds with the transferable name (average cooperation per round 63.0%) than during the rounds with the non-transferable 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 used each group of 6 students as the statistical unit to avoid pseudoreplication) (Fig. 1). In the groups of treatment 2 (PG rounds 11–20 with the non-transferable name and PG rounds 21–25 with the transferable name), the level of cooperation was also significantly higher in rounds with the transferable name (average cooperation per round 66.0%) than in rounds with the non-transferable 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 treatments 1 and 2 depicts a significant overall effect (Fisher 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 transferable name in PG rounds than with the non-transferable 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 transferable name, €1.20 per round with non-transferable name, Wilcoxon signed ranks matched pairs test, z=2.60, P=0.009, n=10, two-tailed). The combined effect is significant (Fisher combination test, chi-square=20.018, P<0.001, df=4).

Discussion

This study shows that the knowledge of being recognized with the same identity (name) in both a public goods game and an indirect reciprocity 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 behaviour could be punished. In non-human animals also,strategic reputation building seems possible. Bshary (2002) suggested that tactical deception in cleaner fish should occur if it pays to alter the optimal behaviour in a situation to induce responses in bystanders (clients), which will produce benefits during future interactions with these bystanders that exceed the momentary costs.

The results of the present study imply that being aware of the fact that any reputation will be transferable can remove the social dilemma from a public goods scenario. This is only a side effect which has, however, impressive consequences. 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 pay-off in this scenario. The pay-off 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), 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. (2002) 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 to the public good, and defectors who join the group but do not contribute to the public good, will coexist through a rock-paper-scissors dynamics even under full anonymity. An experimental study (Semmann et al. 2003) 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 choose to be cooperators, the highest pay-off is achieved as a defector, resulting in an increasing number of players choosing this strategy. When defectors dominate, the highest pay-off can be achieved through choosing to be a loner. Finally, when loners dominate, the public good group size is very small and the highest pay-off 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 is removed, the decisions made in any public goods situation could be recalled in other

social games, and would thereby be connected to reputation building, and 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, the Max-Planck-Institute of Meteorology Hamburg for hospitality, and the anonymous referees for improving the manuscript. The experiments described in this paper were conducted in accordance with local and federal laws in Germany.

References

Alexander RD (1987) The biology of moral systems. de Gruyter, New York

Axelrod R, Hamilton WD (1981) The evolution of cooperation. Science 211:1390–1396

Berkes F, Feeny D, McCay BJ, Acheson JM (1989) The benefits of the commons. Nature 340:91–93

Bolton GE, Katok E, Ockenfels A (2001) University of Wisconsin, Working Paper. http://www.ise.ufl.edu/scmec/katok.pdf

Boyd R, Richerson PJ (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol Sociobiol 13:171–195

Brandt H, Hauert C, Sigmund K (2003) Punishment and reputation in spatial public goods games. Proc R Soc Lond Ser B Biol Sci 270:1099–1104

Broadbent DE (1965) Information processing in the nervous system. Science 150:457–462

Brown SP, Hochberg ME, Grenfell BT (2002) Does multiple infection select for raised virulence? Trends Microbiol 10:401–405

Bshary R (2002) Biting cleaner fish use altruism to deceive image-scoring client reef fish. Proc R Soc Lond Ser B Biol Sci 269:2087–2093

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

Dawes RM (1980) Social dilemmas. Annu Rev Psychol 31:169–193

Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415:137–140

Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. Econ Lett 71:397–404

Fishman MA (2003) Indirect reciprocity among imperfect individuals. J Theor Biol 225:285–292

Gintis H (2000) Strong reciprocity and human sociality. J Theor Biol 206:169–179

Gintis H, Smith EA, Bowles S (2001) Costly signaling and cooperation. J Theor Biol 213:103–119

Hamilton WD (1964) Genetical evolution of social behaviour. J Theor Biol 7:1–52

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 (2002) Volunteering as red queen mechanism for cooperation in public goods games. Science 296:1129–1132

Hoffman E, McCabe K, Smith VL (1996) Social distance and other-regarding behavior in dictator games. Am Econ Rev 86:653–660

Ledyard JO (1995) Public goods: a survey of experimental research. In: Kagel JH, Roth AE (eds) Handbook of experimental economics. Princeton University Press, Princeton, pp 111–194

Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. Proc R Soc Lond Ser B Biol Sci 268:745–753

Lotem A, Fishman MA, Stone L (1999) Evolution of cooperation between individuals. Nature 400:226–227

Milinski M (1990) Information overload and food selection. In: Hughes RN (ed) Behavioural mechanisms of food selection. Springer, Berlin Heidelberg New York, pp 721–737

Milinski M, Semmann D, Bakker TCM, Krambeck HJ (2001) Cooperation through indirect reciprocity: image scoring or standing strategy? Proc R Soc Lond Ser B Biol Sci 268:2495–2501

Milinski M, Semmann D, Krambeck HJ (2002a) Donors to charity gain in both indirect reciprocity and political reputation. Proc R Soc Lond Ser B Biol Sci 269:881–883

Milinski M, Semmann D, Krambeck HJ (2002b) Reputation helps solve the 'tragedy of the commons'. Nature 415:424–426

Mohtashemi M, Mui L (2003) Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism. J Theor Biol 223:523–531

Nowak MA, Sigmund K (1998a) The dynamics of indirect reciprocity. J Theor Biol 194:561–574

Nowak MA, Sigmund K (1998b) Evolution of indirect reciprocity by image scoring. Nature 393:573–577

Ostrom E (1999) Governing the Commons. Cambridge University Press, Cambridge

Ostrom E (2003) Trust & reciprocity: interdisciplinary lessons from experimental research. In: Ostrom E, Walker J (eds) The Russel Sage Foundation Series on Trust, vol VI. Russel Sage Foundation, New York

Parker GA, Chubb JC, Roberts GN, Michaud M, Milinski M (2003) Optimal growth strategies of larval helminths in their intermediate hosts. J Evol Biol 16:362

Pollock G, Dugatkin LA (1992) Reciprocity and the emergence of reputation. J Theor Biol 159:25–37

Rainey PB, Rainey K (2003) Evolution of cooperation and conflict in experimental bacterial populations. Nature 425:72–74

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the ultimatum game. Science 300:1755–1758

Seinen I, Schram A (2001) Tinbergen Institute, Amsterdam Discussion Paper TI2001–003/1. http://netec.mcc.ac.uk/WoPEc/data/Papers/dgruvatin20010003.html, direct http://www.tinbergen.nl/discussionpapers/01003.pdf

Semmann D, Krambeck HJ, Milinski M (2003) Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature 425:390–393

Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. Proc Natl Acad Sci USA 98:10757–10762

Trivers RL (1971) Evolution of reciprocal altruism. Q Rev Biol 46:35

Velicer GJ, Yu YTN (2003) Evolution of novel cooperative swarming in the bacterium *Myxococcus xanthus*. Nature 425:75–78

Wedekind C, Braithwaite VA (2002) The long-term benefits of human generosity in indirect reciprocity. Curr Biol 12:1012–1015

Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. Science 288:850–852