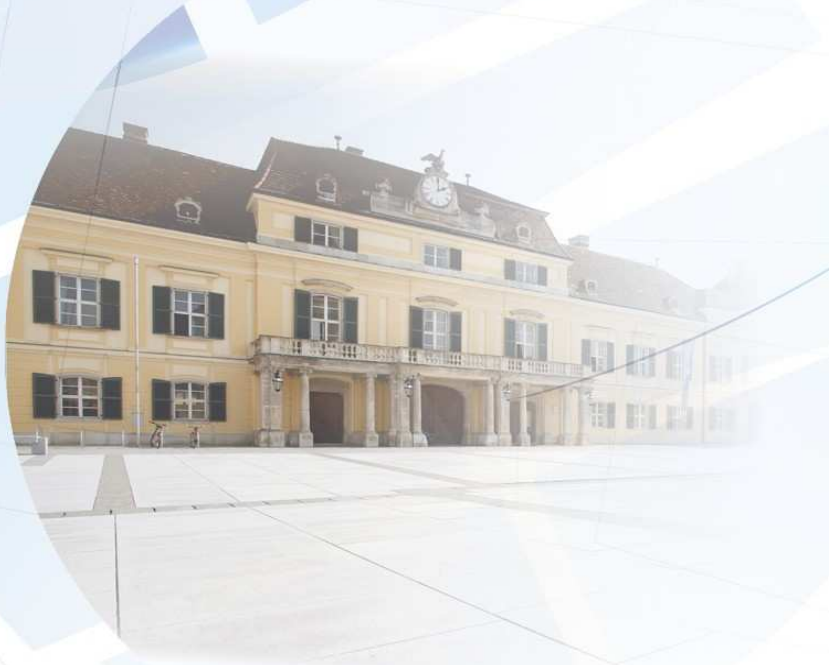
International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

The basis of morality: Richard Alexander on indirect reciprocity

Sigmund, K.

IIASA Interim Report
2013



International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-13-068**

The basis of morality: Richard Alexander on indirect reciprocity

Karl Sigmund (ksigmund@iiasa.ac.at)

Approved by

Ulf Dieckmann
Director, Evolution and Ecology Program

June 2015

The basis of morality

Richard Alexander on Indirect Reciprocity

Karl Sigmund

(Faculty of Mathematics, University of Vienna, and Institute for Applied Systems Analysis, Laxenburg)

(an introduction to an excerpt from Alexander, RD, The Biology of Moral Systems)

In Richard Alexander's *Biology of Moral Systems* (BMS, Alexander 1987), the concept of indirect reciprocity plays a star role. The author firmly states (BMS p.77) that ‚moral systems are systems of indirect reciprocity', and writes (BMS p.95) that ‚systems of indirect reciprocity become automatically what I am here calling moral systems'. One chapter of the book is entitled: ‚Moral systems as systems of reciprocity' (BMS p.93), etc. It should be mentioned in this context that Alexander views 'moral systems' as guides of actions, or standards of conducts, and carefully separates them from the concept of ‚morality', which seems much harder to pin down.

When defining indirect reciprocity, Alexander contrasts it with the simpler concept of direct reciprocity. The latter occurs when ‚the return from the social investment in another individual is expected from the actual recipient of the beneficence… In indirect reciprocity, the return is expected from someone other than the recipient of the beneficence. The return can come from essentially any individual or collection of individuals in the group' (BMS p.85). The concept of indirect reciprocity is also defined in *Darwinism and Human Affairs* (1979), where Alexander writes: 'Reciprocity can be divided into two types. Direct reciprocity occurs when rewards come from the actual recipient of beneficence. Indirect reciprocity, on the other hand, is represented by rewards from society at large, or from others than the actual recipient of beneficence. We engage in both kinds more or less continuously' (p.49).

At the time when Alexander wrote these lines, in the late 'seventies, evolutionary biology was just beginning to acknowledge the importance of direct reciprocity. In particular, Robert Axelrod and William D. Hamilton were using computers to conduct their famous round-robin tournaments of iterated Prisoner's Dilemma games, and to analyse the merits of Tit For Tat, the epitome of reciprocation (Axelrod and Hamilton 1981, see also Axelrod 1984 and Hamilton 1996). The simplest version of a Prisoner's Dilemma

game is obtained if two players can independently decide whether or not to confer a benefit b to their co-player, at cost c to themselves, with 0<c<b. The dominating strategy, here, is to defect, as this maximizes a player's payoff, no matter what the other player does. But if the game is repeated sufficiently often between the same two players, unconditional defection is no good strategy against a Tit For Tat player, i.e., a player who confers a benefit to the co-player in the first round and from then on does whatever the co-player did in the previous round. In particular, Axelrod and Hamilton found that in various computer simulations of repeated Prisoner's Dilemma games, selection led to the emergence of Tit For Tat, and hence to the evolution of cooperation. The work of Axelrod and Hamilton thus confirmed Robert Triver's seminal work (Trivers 1971) which had established reciprocity as the second pillar, next to kin selection, to support altruism in evolutionary biology. In the second edition of Richard Dawkins' *The Selfish Gene*, a chapter was added to celebrate the triumph of Tit For Tat: 'Nice guys finish first' (Dawkins 1989).

Indirect reciprocity is considerably more subtle than direct reciprocity. The latter is based on the principle ‚I'll scratch your back if you scratch mine', whereas the former is based on ‚I'll scratch your back if you scratch someone else's' (Binmore 1994). The merits of this maxim seem less immediate to grasp. They certainly require some sophistication. In Alexander's words, 'indirect reciprocity involves reputation and status, and results in everyone in a social group continually being assessed and reassessed by interactants, past and potential, on the basis of their interactions with others' (BMS p.85). In another statement, 'indirect reciprocity develops because interactions are repeated, or flow among a society's members, and because information about subsequent interactions can be gleaned from observing the reciprocal interactions of others' (BMS p.77). It is this assessment of the actions of others (even if they are not directed at oneself) which is the basis of moral judgements.

Every idea has its fore-runners, and Alexander points out repeatedly that others before him have dealt with generalizations of reciprocity, which he views as 'the binding cement of human social life' (BMS p.111). Alexander was particularly influenced by Trivers (1971, cf. Trivers 1986 and Trivers 2006). Darwin had also anticipated the idea that assessments by others play a fundamental role in human cooperation. In the *Descent of Man*, Darwin wrote that (in contrast to other social animals such as bees or ants), 'man's motive to give aid no longer consists solely of a blind instinctive impulse, but is largely influenced by the praise and blame of his fellow men.' We are all acutely concerned with how we are judged by those around us.

Alexander thus squarely embraced the misanthropic tradition of many thinkers before him, who suspected that costly and seemingly altruistic acts often pay, in the long run, and therefore are not altruistic, even if actors themselves may think so. Their reward, or return, can come from various sources, either individuals, or collections of individuals, or even 'society at large'.

Let us suppose first that the return is provided by individuals. In direct reciprocity, A helps B and B helps A. This yields a net benefit for both. For indirect reciprocity, Alexander mentions two mechanisms, by way of example (BMS p.81). One is of the form: A helps B, B helps C, C helps A. Such cycles of helping were analysed by Boyd and Richerson (1989), who found, however, that they were fragile and unlikely to occur, essentially because the 'return' is so roundabout that the cycle can easily be broken. The other mechanism suggested by Alexander was the following: A helps B; C, observing, later helps A; A helps C. This means that A was rightly judged to be a reliable partner by C, who uses this information to engage with A in direct reciprocity. In this sense, indirect reciprocity acts as a kind of foreplay for direct reciprocity. The corresponding strategy for the repeated Prisoner's Dilemma was termed Observer Tit For Tat (Pollock and Dugatkin 1992): it only deviates from Tit For Tat in the first round, by refusing to help if the co-player, in the last interaction with some third party, has refused to help.

But actually, the system proposed by Alexander works even if direct reciprocity is explicitly excluded. This was shown by a series of models which assumed that no players would ever meet the same co-player again. The principle is: 'A helps B; C, observing, later helps A', which is just as before, except that the appendix 'A helps C', which presumes a second meeting between A and C, is omitted. The continuation in the modified version is implicit: 'D, observing, later helps C', and so on.

It is worth to explore the simplest models of this type (Nowak and Sigmund 1998a,b, Lotem et al, 1999, Nowak and Sigmund, 2005, Pacheco et al. 2006, Sigmund 2010). Thus, suppose that in a large population, two players A and B meet randomly, and each can either provide some help to the other, or refuse to help. (In an equivalent version, one player is randomly assigned the role of potential donor of the help, and the other the role of recipient.) If the same two players never meet again, the Tit For Tat strategy makes no sense. But a closely related variant of discriminating cooperation does. Players using this 'reciprocating' strategy refuse to help those players who

have previously refused to help someone else. In this way, beneficence is channelled towards those players who themselves engage in beneficence. If C observes that A helps B, then C will help A, even if this does not lead to repeated interactions between A and C. In a sense, this is a vicarious return: C returns the help in B's stead.

This model can be made more explicit in various ways. The conditional strategy clearly requires that players have some information about the past behaviour of their co-players. Such information can be incomplete, and still lead to cooperation. It is enough, for instance, to require that (a) with a certain probability q, the potential donor knows whether the potential recipient has refused to help, on some previous occasion, and (b) in the absence of information, the donor is willing to use the 'benefit of doubt'. If the probability q is larger than the cost-to-benefit ratio c/b, cooperation can be sustained in the population: exploiters who never provide help will rarely receive help, and do less well than the discriminating co-operators.

Even in this simple toy-model, it is of paramount importance that players can acquire sufficient information about other group members. Clearly, such 'social scrutinizing' (to use Alexander's term) is facilitated if individuals have a good memory, and spend much of their time together. But it is likely that direct observation is not enough. In all human groupings, individuals exchange information, and communicate what they observe through gossip. Here, the unique language abilities of our species come into play. ('For direct reciprocity, you need a face; for indirect reciprocity, you need a name.' Haigh, personal communication). Conversely, the need to exchange information about others may have been a strong, possibly even the major selective force behind the emergence of the human language instinct (Dunbar 1996).

Both the ability to learn a language and the ability to learn a moral code seem restricted to the human species. Alexander does 'not exclude the possibility that indirect reciprocity […] will eventually be documented in some primates, social canines, felines, cetaceans and some others' (BMS p.85). Only few instances of indirect reciprocation have been documented in non-human species, so far (Bshary and Grutter 2006, Rutte and Taborsky 2007) On the other hand, a large number of economic experiments have shown that humans are highly prone to engage in indirect reciprocity, and that (a) players known to help others usually increase their chances in getting helped by third parties and (b) conversely the propensity to help frequently more than doubles if players know that their decision will be communicated. Help, in this context, is not an altruistic act, but an investment into the social capital of reputation (Wedekind and Milinski

2000, Wedekind and Braithwaite 2001, Seinen and Schram 2001, Milinski et al 2001). This confirms Alexander's view that morality based on indirect reciprocity may be seen as self-serving, because it causes a sufficient number of persons to regard the actor as a good object of social investment (BMS p.109).

The moral assessment of other group members can be captured, in its simplest form, by labelling them 'good' or 'bad', depending on whether they helped or not. Clearly, such a binary assessment leads to a picture in black and white which is much cruder than what we are used to in real life, but it suffices for a proof of principle. Moreover, it raises an intriguing issue. In the most rudimentary form of indirect reciprocity, the discriminating strategy, a cousin of Tit For Tat, extends help to those who are 'good' and refuses help to those who are 'bad'. However, a discriminating player who refuses to help a 'bad' player becomes 'bad' in the eyes of all observers, and therefore less likely to be helped, in turn. In order to maximize the help one receives from discriminators, it would be better to never refuse to help, and thus to stop discriminating. But in a group without discriminators, exploiters go unpunished, and will spread. Cooperation cannot be sustained in the long run.

A remedy coming immediately to mind is to assume that a refusal to help can be justified, if it is directed towards a recipient with a 'bad' image (Nowak and Sigmund 1998a, Leimar and Hammerstein 2001, Panchanathan and Boyd 2003). Such a justified refusal ought therefore not to be labelled as 'bad'. This requires, however, that observers are aware of the reputation of the potential recipient. This, in turn, may require to know the reputation of the recipient's previous recipient, etc. It is questionable whether under normal conditions, individuals have enough information about their group members, or are sufficiently proficient at coping with this information (Milinski et al 2001). Moreover, we still have not described the assessment system completely. While it obviously should be good to refuse help to a bad player, it seems less clear whether giving help to a bad player should be considered as good or bad, for example.

This leads to consider assessment systems which are not only based on whether help is given or not, but also on the reputations of recipient and donor. There are no less than 256 of them, even under the absurdly oversimplified assumption that a player can only be 'good' or 'bad', without intermediate grades (Ohtsuki and Iwasa 2004, Brandt and Sigmund 2004). It turns out that only 8 of them are stable, in the sense that a population which adopts them will cooperate and cannot be invaded by unconditional strategies of always giving or always refusing help (Ohtsuki and Iwasa

2006). The competition of these rudimentary 'moral systems' under conditions which include the possibility of occasional errors in action or judgement turns out to be remarkably difficult to analyse (Ohtsuki and Iwasa 2007,Uchida and Sigmund 2010, Sigmund 2010, Uchida 2011) . Indeed, the status of a given group-member will in general be different for observers using different assessment rules. If additionally, the assessment of a player is based on several actions of that player, or if there are more than two labels for a player's reputation (for instance, 'good', 'bad' and 'indifferent'), the complexity of the moral system explodes. Formalizing ethics appears to be harder than formalizing logic. Practical philosophy defies mathematizing.

It is doubtful whether Alexander would view the investigation of formalized systems of moral assessment rules as useful or relevant for evolutionary biology. Indeed models of indirect reciprocity which isolate it from direct reciprocity, on the one hand, and from group interactions, on the other, are artificial devices, useful for thought experiments but far removed from reality. Alexander is more interested in the fluid, and ever growing boundaries of systems of reciprocity, and the effect of this development on the human psyche. He suggests that 'indirect reciprocity led to the evolution of ever keener abilities to observe and interpret situations with moral overtones' (BMS p.100). 'I regard indirect reciprocity as a consequence of direct reciprocity occurring in the presence of interested audiences – groups of individuals who continually evaluate the members of their society as possible future interactants from whom they would like to gain more than they lose' (BMS p.93). In particular, 'we use motivation and honesty in one circumstance to predict actions in others...Humans tend to decide that a person is either moral or not, as opposed to being moral in one context and immoral in another' (BMS p.94). And indeed, it is remarkable that ancient philosophers saw virtue as the attribute of a person, whereas contemporary philosophers speak of virtue as the attribute of an act, or a decision.

There seem to exist mechanisms of indirect reciprocity which are not based on reputation. For instance, we could consider a situation where first, A helps B, and then, B helps a third party C. There are many examples of this. If someone holds the door open to us, we are likely to hold open the door for the next. This type of indirect reciprocity, where the recipient returns the help, but not to the actual donor, has also been observed in economic experiments and seems less easy to explain theoretically.

Moreover, negative interactions may also be reciprocated. Alexander usually speaks of 'rewards' being returned, but retaliation clearly is also a wide-spread form of reciprocation, and he mentions it, for instance on a

table describing various forms of indirect reciprocity (BMS p.86), or by stating that 'systems of indirect reciprocity involve promises of punishment as well as reward' (BMS p.96). In the last two decades, experimental economists have uncovered a wide-spread propensity to punish those who are perceived as cheaters. So-called 'peer-punishment' is very effective at promoting cooperation (Fehr and Gächter 2000). Interestingly, many individuals are ready to incur personal costs to punish norm-breakers, even if they themselves were not affected by the misdeameanour, but merely observed it. This is certainly also a form of indirect reciprocity.

So far, we have considered reciprocity based on returns by individuals. Alexander stressed on several occasions that such a return could also be provided by collections of individuals, as when he writes 'indirect reciprocity, whereby society as a whole or a large part of it provides the reward for the beneficence' (BMS p.105). Societies provide not only rewards for beneficence, but also punishment for free-riding. In fact, most punishment is not meted out by irate individuals in the form of 'peer-punishment', but rather by institutions (Ostrom 1990; Yamagishi 1986; Sigmund et al. 2010). Institutions can be viewed as tools enabling communities to provide positive or negative incentives. There is a striking similarity between institutionalized punishment of free-riding and the repression of competition encountered in many examples of cooperative groupings encountered in biological evolution (Frank, 1995).

In a particularly interesting aside on indirect reciprocity, Alexander mentions that the reward which an individual obtains for acts of beneficence can simply consist in the success of the group (BMS p.94). This relates to an aspect which is crucial for the role played by moral systems, in Alexander's view.

Indeed, following Keith (1947), who stressed the competition between groups as a main factor shaping human evolution, Alexander sees 'morality as a within-group cooperativeness in the context of between-group competition' (BMS p.153). According to his celebrated 'balance of power'-argument, the often lethal competition between families, bands, tribes and nation was the chief selective force shaping human evolution. 'In no other species do social groups have as their main jeopardy other social groups of the same species' (BMS p.77).

Alexander accordingly writes (BMS p.194) that 'indirect reciprocity is more complex than is usually realized partly because of long-term benefits from being viewed as an altruist, and partly because one must take into account benefits to the individual that accrue from the success of his group in

competition with other groups.' This latter aspect is not, in general, associated with 'indirect reciprocity' nowadays. The former aspect has monopolized the meaning. The 'long-term benefits for being viewed as an altruist' help to establish generalized exchange systems working to mutual advantage, even in the absence of strife between groups.

The role of group selection has been hotly contended by evolutionary biologists during the last half-century. Mathematical models often use alternative expressions, such as kin selection, or multi-level selection, sometimes with the intention of avoiding semantic quarrels, and usually with the result of exacerbating them. But there seems no reason to avoid the name 'group selection' when one speaks of groups fighting and annihilating each other. Individuals have to balance, in such contests, their well-being within the group with the well-being of their group.

This viewpoint was shared by Darwin, who wrote: 'There can be no doubt that a tribe including many members who [...] were always ready to give aid to each other and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection.' He did not say: '...and this is group selection', but he obviously was aware of the tension between individual and group selection when he wrote: 'He who was ready to sacrifice his life [...] would often leave no offspring to inherit his noble nature. Therefore it seems scarcely possible (bearing in mind that we are not here speaking of one tribe being victorious over another) that the number of men gifted with such virtues could be increased through natural selection.' The term in parentheses clearly indicates that Darwin saw no way of explaining the evolution of such self-sacrificing traits other than by violent inter-group conflict. In another passage, Darwin stressed that 'extinction follows chiefly from the competition of tribe with tribe, and race with race.' To many ears, today, this sounds politically incorrect. But Darwin was very conscious of this dark side of human nature. Some of the most remarkable examples of human cooperation occur in war, and other forms of lethal conflict between groups, and it is well-known that day-to-day solidarity dramatically increases in societies threatened from the outside.

That group selection can favour the emergence of cooperative traits has been shown in countless models. It is all the more remarkable that the models and experiments on indirect reciprocity which have been mentioned so far assume a single, well-mixed population. They show that indirect reciprocity can work even if the population is not structured into competing groups. Individuals who deviate from the cooperative norm will

have, on average, their long-term payoff reduced, and thus are unlikely to be copied by others. This confirms Alexander's view that it is fallacious to assume that morality inevitably involves some self-sacrifice (BMS p.161). However, it confirms it in a set-up which is different from Alexander's.

A final remark: economists have also been investigating extensions of the concept of reciprocity, in parallel to evolutionary biologists. Their point of departure was usually the so-called 'folk theorem on repeated games', which states that if the probability of another round between the same two players is sufficiently high, cooperation can be sustained by so-called trigger strategies, of which Tit For Tat is but one example. A rational player would forego the exploitation of a co-player in one round if this jeopardizes collaboration in all future rounds. It is clear that if players interact only once, a cheater cannot be held to account by its victim, and therefore personal enforcement must be replaced by community enforcement. Game theorist have shown that even if information is transmitted only imperfectly, cooperation can be sustained, based on trigger strategies adopted by the whole community. No rational player has an interest in deviating unilaterally (Rosenthal 1979, Sugden 1986, Kandori 1992, Ellison 1994, Okuno-Fujiwara and Postlewaite 1995, Bolton et al 2004a). The interest of economists in indirect reciprocity has been singularly heightened, in recent years, by the fact that one-shot interactions between anonymous partners become increasingly frequent in today's society. Web-based auctions and other forms of e-commerce occur between strangers who never meet face to face. They are built on rudimentary reputation mechanisms similar to those which were developed in the simplest formal models capturing Richard Alexander's ideas on indirect reciprocity (Bolton et al 2004b, Keser 2003).

Richard Alexander recognized indirect reciprocity based on reputation and status as major factor in the emergence of moral systems in human societies. It is amazing that the same simple, robust mechanisms that shaped early hominid societies now play a central role in shaping the internet civilisation.

References

Alexander, R.D. 1979. Darwinism and Human Affairs, Seattle, Univ. Washington Press.

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

Axelrod, R. 1984. The Evolution of Cooperation, Basic Books, New York (reprinted 1989 in Penguin, Harmondsworth).

Axelrod, R, and Hamilton, WD. 1981. The evolution of cooperation, Science 211:1390-1396.

Binmore, K. 1994. Playing Fair: Game Theory and the Social Contract, Cambridge, MA: MIT Press.

Bolton, G., Katok, E., Ockenfels, A. 2004a. Cooperation among strangers with limited information about reputation, Journ. Pub. Econ. 89:1457-1468.

Bolton, G., Katok, E. and Ockenfels, A. 2004b. How effective are online reputation mechanisms? An experimental investigation, Management Science, 50:1587-1602.

Boyd, R and Richerson, P.J. 1989. The evolution of indirect reciprocity, Social Networks 11:213-236.

Brandt, H. and Sigmund, K. 2004. The logic of reprobation: assessment and action rules for indirect reciprocity, Journ Theor. Biol. 231:475-486.

Bshary, R. and Grutter, A.S. 2006. Image scoring causes cooperation in a cleaning mutualism, Nature 441:975-978.

Dawkins, R. 1989. The selfish gene, 2nd edition, Oxford: Oxford Univ. Press.

Dunbar R. 1996. Grooming, gossip and the evolution of language, Harvard Univ. Press.

Ellison, G. 1994. Cooperation in the Prisoner's Dilemma with anonymous random matching, Review of Economic Studies, 61:567-588.

Fehr, E. and Gächter, S. 2002. Altruistic punishment in humans, Nature 425:785-791 .

Frank, S.A. 1995. Mutual policing and the repression of competition in the evolution of cooperative groups. Nature 377:520-522.

Hamilton, W.D. 1996. Narrow Roads of Gene Land, Vol I, Freeman, New York.

Kandori, M. 1992. Social norms and community enforcement, The Review of Economic Studies 59:63-80.

Keith, A. 1947. Evolution and Ethics, Putnam and Sons, New York.

Keser, C. 2002. Trust and Reputation Building in e-Commerce, IBM Systems Journal 42:498-506.

Leimar, O. and Hammerstein, P. 2001. Evolution of cooperation through indirect reciprocation, Proc R Soc Lond B, 268:745-753.

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

Milinski, M., Semmann, D., Bakker, T.C.M. and Krambeck, H. J. 2001. Cooperation through indirect reciprocity: image scoring or standing strategy? Proc Roy Soc Lond B 268:2495-2501.

Nowak, M.A. and Sigmund, K. 1998a. Evolution of indirect reciprocity by image scoring, Nature 282:462-466.

Nowak, M.A. and Sigmund, K. 1998b. The dynamics of indirect reciprocity, Journ Theor. Biol. 194:561-574.

Nowak, M.A. and Sigmund, K. 2005. Evolution of indirect reciprocity, Nature 437:1291-1298.

Ohtsuki, H. and Iwasa, Y. 2004. How should we define goodness? Reputation dynamics in indirect reciprocity, Journ Theor. Biol. 231:107-120.

Ohtsuki, H and Iwasa, Y. 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity, Journ Theor. Biol. 239:435-444.

Ohtsuki, H and Iwasa, Y. 2007. Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation, Journ Theor. Biol. 244:518-531.

Okuno-Fujiwara, M and Postlewaite, A. 1995. Social
Norms in Matching Games, Games and Economic Behaviour, 9:79-109.

Ostrom, E. 1990. Governing the Commons. Cambridge: Cambridge Univ Press.

Pacheco, J, Santos, F and Chalub, F. 2006. Stern-judging: a simple, successful norm which promotes cooperation under indirect reciprocity, PLOS Computational Biology 2:e178.

Panchanathan, K. and Boyd, R. 2003. A tale of two defectors: the importance of standing for evolution of indirect reciprocity, Journ Theor. Biol. 224:115-126.

Pollock, G.B. and Dugatkin L.A. 1992. Reciprocity and the evolution of reputation, Journ Theor. Biol. 159:25-37.

Rosenthal, R.W. 1979. Sequences of games with varying opponents, Econometrica 47:1353-1366.

Rutte, C. and Taborsky, M. 2007. Generalized reciprocity in rats. PLoS Biology 5:1421-1425.

Seinen, I. and Schram, A. 2001. Social status and group norms: indirect reciprocity in a helping experiment, European Econ. Review 50: 581-602.

Sigmund, K. 2010. The Calculus of Selfishness, Princeton, NJ: Princeton Univ. Press.

Sigmund, K., De Silva, H. Traulsen, A. and Hauert, C. 2010. Social learning promotes institutions for governing the commons, Nature 466:861-863.

Sugden, R. 1986. The Economics of Rights, Cooperation and Welfare, Oxford: Basil Blackwell.

Trivers, R. 1971. The evolution of reciprocal altruism, Quart Rev Biol 46:35-57.

Trivers, R. 1986. Social Evolution, Menlo Park, CA: Benjamin Cummings.

Trivers, R. 2006. Reciprocal altruism: 30 years later, in Cooperation in Primates and Humans: Mechanisms and Evolution, ed P.M. Kappeller and C.P. van Schaik, 67-83, Berlin, Springer.

Uchida, S. 2010. Effect of private information on indirect reciprocity, Phys. Rev. E 82:doi 10.1103/PhysRevE.82.036111.

Uchida, S. and Sigmund, K. 2010. The competition of assessment rules for indirect reciprocity, Journ Theor. Biol. 263: 13-19.

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

Wedekind, C. and Braithwaite, V.A. 2002. The long-term benefits of human generosity in indirect reciprocity, Curr Biol. 12:1012-1015.

Yamagishi, T. 1986. The provision of a sanctioning system a a public good. Journal of Personality and Social Psychology 51:110-116.