Journal of Institutional Economics (2014), 10: 2, 197–230
© Millennium Economics Ltd 2013 doi:10.1017/S1744137413000374
First published online 23 December 2013

Eleven mechanisms for the evolution of cooperation

MICHAEL A. ZAGGL*
TUM School of Management, Technische Universität München, Arcisstr. 21, 80333, Munich, Germany

Abstract: Cooperation is one of the basic elements of social life. It is essential for emergent social phenomena, such as the formation of families, groups, and societies. However, evolutionary forces counter cooperation. The trait of supporting others is dominated by selfish behavior. In the last few decades scientists, in particular biologists, achieved extraordinary progress regarding the question of how cooperation is possible despite of evolutionary forces. This produced an enormous amount of literature. This paper identifies and reviews the known solutions explaining cooperation under evolutionary forces. Using bibliometric methods in combination with extant review articles and traditional reviewing of original literature, it is possible to isolate 11 mechanisms of cooperation under the conditions of evolution. Developing a categorization of the mechanisms according to shared characteristics establishes a fundamental framework for institutional and mechanism design activities. Implications for future research paths are identified, in particular for the mechanism of indirect reciprocity.

1. Introduction

Cooperation is the glue that binds individuals together and allows for the emergence of social structures on higher levels, such as families, groups, organizations, nations, and civilizations. Our understanding of humankind is impossible in the absence of cooperation. Cooperation is thus an elementary force for the disciplines of the social sciences. However, the most consistent body of theory addressing cooperation is not allocated in one of the disciplines of the social sciences. Instead, it can be found in biology (e.g., Nowak and May, 1992; Nowak and Sigmund, 1998a, 2005; Nowak *et al.*, 2004; Ohtsuki *et al.*, 2006; Trivers, 1971). The strong relation between cooperation and the theory of evolution may be the reason for the dominance of biology. Overall, a large amount of literature has developed in this field. However, as an indispensable element of sociality, cooperation plays a central role particularly in the social sciences (e.g., Axelrod, 1984; Axelrod and Hamilton, 1981), especially for economics (e.g., Fehr *et al.*, 1997, 1998; Ockenfels, 1993). Here, the mechanisms

*Email: michael.zaggl@tum.de

of cooperation are a potent part for the design of organizations, solving public good problems, and designing incentive systems – just to name a few examples. Notably, the mechanisms are fundamentally important for modern economic theories, in particular for fairness theory (Bolton and Ockenfels, 2000; Dufwenberg and Kirchsteiger, 2004; Fehr and Schmidt, 1999; Rabin, 1993).

From an abstract perspective, the mechanisms of cooperation play the same role for institutional or mechanism design that natural laws play for engineering. This picture of the relation between natural sciences and engineering can be quite useful to argue the scope of the mechanisms of cooperation. By perceiving the mechanisms of cooperation as natural laws, institutional design becomes the corresponding engineering science – i.e., the usage of the natural laws for certain purposes. Engineering is always dependent on the limits and possibilities determined by natural laws. Analogously, institutional design persistently depends on the mechanisms of cooperation.

Because of the large amount of literature addressing cooperation, this study aims to collect and review the extant mechanisms of cooperation. This task is approached from the perspective of economics and institutional design/policy making. For that purpose, the mechanisms are organized in a framework. The insights resulting from this endeavor enrich the perspective on the issue of cooperation for the social scientist. It extends the set of tools for designing systems that aim the increase of benefits from social interactions. For example, the challenge of efficient knowledge sharing inside organizations is a highly relevant topic (e.g., Kogut and Zander, 1996; von Krogh *et al.*, 2012), which can be approached with detailed knowledge of the mechanisms of cooperation. Furthermore, the same knowledge about the mechanisms is a precondition for the reduction of undesired cooperation, such as collusion.

Three approaches for the identification of the mechanisms of cooperation are conducted. First, a bibliometric analysis is used to identify structures inside the literature corresponding with mechanisms of cooperation. Second, extant review articles that attempt to collect mechanisms of cooperation are used. Third, original literature is consulted.

The remainder of this paper is structured as follows. Next, the method of co-citation analysis as well as the data set and its results are introduced. In Section 3, the identified mechanisms of cooperation are introduced and in Section 4 bibliometric data is used to observe the mechanisms' impact in the literature over time. In Section 5, a framework is developed that structures the mechanisms. The paper closes with a conclusion in Section 6.

2. Co-citation analysis for identifying mechanisms of cooperation

For the identification of mechanisms of cooperation inside the existing literature, the method of co-citation analysis is employed. It is an established method for analyzing the structures inside scientific fields (e.g., Osareh, 1996). In general, the method of co-citation is suitable for identifying the relational structure between publications. One of its central characteristics is external assignment: The relationship between two publications is not created through the citation of the first publication by the later, but through their combined citation in a third publication. Consequently, there is no influence of the considered publications. Instead, the third publication determines the relation. This diminishes the influence of strategic citing. The idea of co-citation analysis is that the frequent citation of two entities of literature in combination with several publications is an indicator for shared or similar content. Structuring all such relations in a network enables drawing conclusions about gravity centers, subfields, and foci of a research field (Chen *et al.*, 2010; Jarneving, 2005; Small, 1980).

Data set description

The data set used for the co-citation analysis is collected from the bibliometric database *ISI Web of Science*, which is part of *Thomson Reuters' Web of Knowledge*. The majority of the publications available in these databases are natural science and social science journal articles from (according to Thomson Reuters) high-quality journals. The data set is divided in two separate sets. The first set addresses the field of evolutionary game theory. This field is chosen because it can be seen as being strongly related to cooperation. The methodology of evolutionary game theory is often used for approaching the phenomenon of cooperation (e.g., Maynard Smith, 1982, 1991). The second data set addresses the evolution of cooperation. It is expected that the topic of the evolution of cooperation directly addresses mechanisms of cooperation.

For collecting the first data set, the ISI Web of Science has been queried for the keywords 'evolutionary game theory'. Redundancies were found and erased. The remaining 930 publications are the data set for the bibliometric analysis. In addition, checking for errors inside the data, especially in the reference lists of the publications, is necessary because data from the ISI Web of Science contains errors, such as redundant entries in bibliographies and misspelled author names.

Instead of focusing on the 930 publications from the ISI Web of Science query, the references of these publications are taken into account. This procedure provides a more general data set because publications in the form of books, conference proceedings, and the like are systematically underrepresented in the ISI Web of Science. In the form of citation, however, all kinds of publications are possible. After cleansing, the 930 articles cite in total 38,777 references. Thus, the average article cites 41.7 sources. To conduct the co-citation analysis, the data set must be restricted to focus on the most influential references. Therefore, only entities are selected that are cited at least 10 times. To weight the relations in the co-citation network, the CoCit-value is calculated (Gmür, 2003) and an optimal threshold is determined by maximizing its discriminatory power. Altogether, the resulting network consists of 243 nodes represented by the publications and 559

Table 1. Overview about clusters of evolutionary game theory

	Label	Nodes	Links	Density
1	Spatiality in cooperation	95	258	0.058
1.1	Spatial influences on cooperation	32	72	0.145
1.2	Graph selection theory	29	110	0.271
1.3	Evolutionary games in finite populations	21	34	0.162
1.4	Optional participation	6	6	0.4
1.5	Evolutionary stable strategies in finite populations	4	5	0.833
1.6	Empirics of networks	3	3	1.0
2	Ecological equilibria concepts	9	15	0.416
3	Costly signaling	8	18	0.643
4	Continuous stable strategies and long-term evolution	8	10	0.357
5	Coordination and stochastic conditions	8	8	0.286
6	Costly punishment in humans	7	7	0.333
7	Evolutionary equilibria concepts for game theory	6	5	0.333
8	Evolutionary dynamics and branching	4	5	0.833
9	Empirical observations of direct reciprocity in animals	3	3	1.0

edges represented by the links that are at least as strong as the threshold of the CoCit-value. The network is subdivided into nine components.

The second data set is collected in the same way as the first one. The ISI Web of Science has been queried, this time for the keywords 'evolution of cooperation'. The resulting data set consists of 3,314 records after cleansing and these records cite 150,546 references. All entities that are cited at least 15 times are taken into account. Using the optimal CoCit-value threshold, the resulting network consists of 15 components made up of 182 publication nodes connected by 341 links.

Co-citation analysis results

Applying the Newman algorithm for identifying communities inside networks (Clauset *et al.*, 2004) to the nine components of the co-citation analysis of evolutionary game theory subdivides the largest component, hereinafter referred to as *cluster*, into six communities, hereinafter referred to as *subcluster*. Table 1 lists all clusters' and subclusters' labels with its number of nodes, number of links, and its network density. As expected, some of the (sub)clusters of the co-citation network of evolutionary game theory address mechanisms of cooperation. The largest cluster is labeled 'Spatiality in cooperation'; its content is about the influence of spatial conditions on cooperation. One of its subclusters (Subcluster 1.2) addresses graph selection. Graph selection is a mechanism of cooperation. It addresses the influence of spatiality represented as graphs on cooperation. Another subcluster (Subcluster 1.4) addresses the topic of voluntary or optional participation. Its publication nodes show that the possibility for people to receive a small payoff when not participating in a public good game (i.e., neither contributing nor receiving its benefits) can elicit cooperation. Thus, the subcluster's content represents a mechanism of cooperation. Cluster 3 addresses

Table 2. Overview about clusters of evolution of cooperation

	Label	Nodes	Links	Density
I	Spatial cooperation	56	152	0.099
II	Kinship selection/inclusive fitness theory	20	29	0.153
III	Cooperation and exploitation by parasites	16	19	0.158
IV	Indirect reciprocity	15	30	0.286
V	Skew theory (reproductive partitioning) particularly in social insects	14	21	0.231
VI	Direct reciprocity and tit-for-tat in fishes (particularly sticklebacks and guppies)	12	26	0.394
VII	Cooperation between organizations	12	19	0.288
VIII	Cooperative structuring of unicellular organisms to create multicellular structures	7	7	0.333
IX	Cooperation in humans particularly by food sharing	6	9	0.6
X	Cooperation on the example for *Rhizobium* and nitrogen fixation	6	8	0.533
XI	Religion as costly signaling	4	6	1.0
XII	Reciprocity and cheating in cleaner fishes	4	6	1.0
XIII	Evolutionary stability in the repeated prisoner's dilemma	4	3	0.5
XIV	Optional participation	3	3	1.0
XV	Cooperation in chimpanzees	3	3	1.0

the mechanism of cooperation called costly signaling or the handicap principle. A handicap is used to demonstrate extraordinary fitness or quality. This induces cooperation with a partner who can interpret the signal. Another mechanism of cooperation is costly punishment (Cluster 6). Spending resources for punishing uncooperative behavior makes cooperation more attractive. Finally, Cluster 9 addresses reciprocal behavior in the context of animals. Making the own behavior conditional to opponent's past behavior can elicit cooperation.

The co-citation analysis of the evolution of cooperation results in a network structure consisting of 15 disjoint clusters. The clusters are numbered by roman numbers (I–XV) to avoid confusion with the clusters from the co-citation analysis of evolutionary game theory. The Newman algorithm for identifying communities (Clauset *et al.*, 2004) does not lead to a subdivision of the clusters into subcluster. Table 2 provides an overview.

The examination of the evolution of cooperation reveals some clusters addressing mechanisms of cooperation. Cluster I in the co-citation network of the evolution of cooperation is very similar to Cluster 1 of the co-citation network of evolutionary game theory. Cooperation in the context of spatiality is the topic here as well. The second largest cluster (Cluster II) of the co-citation network of the evolution of cooperation addresses kin selection or inclusive fitness theory, which is a mechanism of cooperation explaining mutual and unilateral support between relatives. Cluster IV addresses indirect reciprocity, which can explain cooperation in asymmetric and unrepeated interactions. Cluster V focuses on reproductive skew theory; thus, its literature attempts to explain the disparity among group members' reproduction rates. This content is related to the mechanism of group selection. As in the co-citation analysis of evolutionary game theory, direct reciprocity occupies a cluster in the co-citation analysis of the evolution of cooperation (Cluster VI). Here, reciprocity is also examined based on empirical studies of animals, but it is focused on fish species. Cluster XI addresses costly signaling which is a mechanism of cooperation already identified in the co-citation analysis of evolutionary game theory. Optional participation is represented by Cluster XIV. Some other clusters of the co-citation network of the evolution of cooperation address cooperation (e.g., Clusters VII, X, and XV), but they do not focus on particular mechanisms.

Overall, the co-citation analyses of evolutionary game theory and the evolution of cooperation reveal a considerable number of mechanisms of cooperation. Several mechanisms are addressed in both analyses. The mechanism of optional participation is not addressed in any of the extant review articles and was only identified because of the co-citation analyses. Furthermore, costly signaling has not been addressed by the extant review articles. The bibliometric examination shows in addition that the issue of cooperation in spatial structures is a particularly dynamic and current topic.

3. Collection of the mechanisms of cooperation

This section introduces the mechanisms of cooperation. To this end, the identified mechanisms from the co-citation analysis are complemented by original literature and extant review articles. The review articles are Bergmüller *et al.* (2007), Dugatkin (1997b), Lehmann and Keller (2006), Nowak (2006), Stevens *et al.* (2005), Taylor and Nowak (2009), and West *et al.* (2006). They are chosen because they already attempted to collect mechanisms of cooperation.

Kin selection

One approach for explaining cooperation is to consider the genetic relatedness of the interacting individuals. Individual life can only survive for a limited time and therefore the only way to conserve genes is copying. Relatives, in particular offspring, are containers of the genes. Thus, the genes can survive independently from the carrier. Supporting the offspring increases the genes' *inclusive fitness*.

However, kin selection does not lead to stability of a blind support by the replicators to their copies. Through sexual propagation and mutation, the degree of genetic material shared between replicator and copy varies. Consequently, the strength of the relatedness between the supporting and the supported individual in a relationship of kin selection must be taken into account. Furthermore, the cost for the supporter, in addition to the benefit for the supported, is crucial. The relation between these parameters has been discussed early. John B. S. Haldane is regarded as the first who discovered the theory about inclusive fitness or kin selection. The formalized theory of kinship is the legacy of William

D. Hamilton, who developed the rule for the stability of inclusive fitness. The genetic relatedness (*r*) of the players must exceed the ratio of the costs (*c*) for the supporter and the benefit for the supported (*b*) to lead to cooperation (Hamilton, 1964a, 1964b): $c/b < r$.

Possibilities to observe kin selection in nature are numerous. For instance, the eusocial behavior of the naked mole-rat is characterized by the support of the queen's offspring by her sisters instead of breeding themselves (Jarvis, 1981). This can be explained by their strong genetic relatedness, which is in an average colony higher than 80% (Reeve *et al.*, 1990). The most apparent example for kin selection is the colony life of social insects,[1] which can be found in ants, wasps, and honeybees.[2]

Empirical evidence for kin selection in humans can be taken from the area of evolutionary psychology. For example does care giving by grandmothers of their grandchildren depend on the fact whether she is the grandchildren's mother's mother or its father's mother. The maternal grandmother provides significantly more intense support to her grandchildren than the paternal grandmother does (Euler and Weitzel, 1996). This is because through females the parenthood is confident but through males it is not. Therefore, costly care taking is directed to the one with the higher probability of genetic relatedness. Examples for kin selection in economics can be found in all cases of nepotism (e.g., Laker and Williams, 2003). Moreover, the strength of family businesses relies on cooperation caused by kin selection.

Kin selection is represented in the co-citation analysis of the evolution of cooperation (Cluster II) and in all review articles. The co-citation analysis reveals that kin selection plays a considerable role in the literature because it is the second-largest cluster in the analysis of the evolution of cooperation. The main publication outlet is the *Journal of Theoretical Biology*, which contains more than 21% of the publications in the cluster. The average article is published between 2000 and 2001. This indicates some activities in the topic despite kin selection theory's age.

Green beard selection

The reasoning about the mechanism with the peculiar name *green beard selection* is straightforwardly derived from kin selection. Whereas in kin selection relatives are supported because they have partly the same genes, the same happens in green beard selection but without the necessity of kinship. When genes in different

1 It can be seen as irony that Darwin (1872: 292) considered the existence of social insects as a serious difficulty for natural selection theory. In the light of kin selection, however, the existence of social insects even can be seen as support for his theory.

2 They are characterized by haplodiploidy. This causes a close relatedness of the females to their sisters, on average to 75% (Andersson, 1984: 166). However, haplodiploidy seems to be no precondition for eusociality. Termites are not characterized by a strong genetic relatedness, but they live eusocial (Thorne, 1997). For an approach to align these facts, see Nowak *et al.* (2010).

organisms are identical, support can be evolutionary stable. However, there is the need for a trait that can be recognized by potential supporters for the purpose of identifying carriers of the particular gene. Therefore, this trait must be a characteristic of the phenotype. This trait can be anything conspicuous, thus also a green beard.[3]

The term 'green beard' coins Dawkins (2006 [1976]). He derives green beard selection theoretically but argues that the likelihood of green beard selection occurring is low. One reason for that is the low probability for two green beards meeting each other. In kin selection, this is no problem because here the proximity is given naturally. For green beard selection, however, this must be alleviated by a high degree of mobility. An additional problem is that the same gene that is responsible for the green beard must also determine the degree of support for the other green beards, because otherwise opportunity is given for free-riders by adapting the mimicry of the identifying trait but not the tendency for cooperation. These free-riders can therefore benefit from the cooperation by the true green beards, but save the costs for supporting others. According to this, the strategy of free-riding will invade and displace the true green beards.

By assuming that cheating is not possible, Hamilton's rule applies as much for green beard selection as for kin selection. By weakening this assumption, the possibility of cheating induces that the relatedness (r) must be increased by the inverse of the probability (p) of meeting a cheater: $c/b < r \cdot (1 - p)^{-1}$.

In spite of green beards' low likelihood for occurring, empirical research seems to show that it can evolve in animals (e.g., Keller and Ross, 1998). Much more information based on scientific literature in context of green beard selection is provided for humans. Genetic similarity theory (Rushton *et al.*, 1984) avers that genetically similar individuals are able to detect each other. Ethnic nepotism is just one salient example. In this context, it should be mentioned that green beard selection has the potential for contributing to social identity theory (Tajfel and Turner, 1986) from an evolutionary viewpoint.

Direct reciprocity

Direct reciprocity enables cooperation even between unrelated individuals. Individuals that for any reason trust in each other's reciprocation of an action that is costly for the actor but advantageous for the opponent can end up in mutual beneficial cooperation. Direct reciprocity depends on repeated interactions. Under this condition, and when the players react sensitively to each other's actions, the folk theorem becomes relevant (see, e.g., Aumann, 1985; Rubinstein, 1979) and makes cooperation a solution in Nash equilibrium. Indeed, under these conditions, cooperators perform better than defectors.

3 Only one review article addresses green beard selection (Lehmann and Keller, 2006).

The well-known computer tournaments by Axelrod and Hamilton (1981) and Axelrod (1984) show that in the prisoner's dilemma, the 'shadow of the future' is the crucial element for the evolution of cooperation in direct reciprocity. It points out that future interactions influence current decisions. Formally, it is the probability that the game is repeated in the next iteration. It works as a discount rate of future payoffs. A general rule for the cooperation under direct reciprocity is based on the shadow of the future (ω): $c/b < \omega$ (Nowak, 2006).

To profit from repeated interactions, players need a memory at least capable of storing the result from the last interaction. Another requirement of direct reciprocity is that noise, in the meaning of misunderstandings of the opponents' actions, must not be too strong (e.g., Molander, 1985; Nowak and Sigmund, 1993).

In humans, direct reciprocity seems to be strongly related to emotions. Most importantly, the emotion of gratitude plays a central role for reciprocity. From an evolutionary viewpoint, one could even argue that gratitude and also revenge are results of adaptation processes to enable direct reciprocity.

From an analytical perspective, direct reciprocity has been examined first by Trivers (1971). He analyzed reciprocal behavior based on cleaning fishes and cleaning shrimps together with their symbiotic partners, the predator fishes. After a cleaner had served by removing parasites and leftovers from the mouth of a predator fish, the latter could take the opportunity and eat the cleaner. This, however, does not happen normally. About 10 years later, the previously mentioned works (e.g., Axelrod, 1984; Axelrod and Hamilton, 1981) focus on the strategies of participants in repeated games. Therefore, computer-based simulations have been combined with genetic algorithms, and earlier, computer tournaments have been conducted. The result of identifying tit-for-tat as the best-performing strategy is well known. Later in the history of research about direct reciprocity, the assumption of absence of noise has been weakened (e.g., Nowak and Sigmund, 1993). In consequence, other strategies could demonstrate their advantages (e.g., win-stay-lose-shift).

Direct reciprocity is represented in both co-citation analyses (Cluster 9 and Cluster VI) and in all of the review articles. This reflects not only its importance for cooperation, but also its role in the scientific development of cooperation theory. However, the clusters in the co-citation analyses are not very prominent. Their focus is shaped by the discipline of biology. The main publication outlets are *Behavioral Ecology and Sociobiology* (Cluster VI) and *Nature* (Cluster 9); and the average publication is from 1989, 1990 respectively.

Overall, direct reciprocity provides an excellent explanation for the evolution of cooperation, particularly in humans. The ability to memorize is well developed in humans and seems to be a vital element in combination with reciprocity for the origin of societies. The mechanism's potential for cooperation between unrelated individuals makes reciprocity to be one of the most important aspects for human interactions in general.

Indirect reciprocity

Indirect reciprocity, in stark contrast to direct reciprocity, is not based on repeated interactions between the same individuals (e.g., Boyd and Richerson, 1989; Nowak and Sigmund, 1998a, 1998b; Sigmund, 2012). Thus, it provides a possible explanation for cooperation between individuals that have never met before and the likelihood of meeting again is negligible. Whereas direct reciprocity can be simplified to '*I help you and you help me*', the case of indirect reciprocity works by '*I help you and somebody else helps me*' (Nowak and Sigmund, 2005: 1,291; Taylor and Nowak, 2009: 44, 45). This opens a potential that is much larger for welfare increase than the potential of direct reciprocity. The reason is that indirect reciprocity is not limited to the same pair of individuals. Instead, other links with larger benefit-to-cost ratios can be utilized. Therefore, indirect reciprocity can be seen as a generalization of reciprocity and it allows the exchange of resources between individuals that cannot pay their opponent directly.

Existing models and research activities regarding indirect reciprocity are methodologically based on mathematics (e.g., Nowak and Sigmund, 1998a, 1998b; Ohtsuki and Iwasa, 2004; Sigmund, 2010), on computer-based simulation (e.g., Brandt and Sigmund, 2004; Leimar and Hammerstein, 2001; Nowak and Sigmund, 2005; Ohtsuki *et al.*, 2009), and, recently, on human subject experiments (e.g., Engelmann and Fischbacher, 2009; Milinski *et al.*, 2001, 2002).

In general, indirect reciprocity seems to become more relevant because of an increase in specialization. The more specialization, the more diversified are the consumption possibilities, and the less frequent are the repeated interactions between the same individuals in relation to the total number of interactions. For instance, an individual who needs dental treatment and a new garden gate sees for both issues the blacksmith when she or he is part of a rather low-specialized society. Under high specialization, however, she or he has to interact with two completely different suppliers, but each time only once. In this environment, a kind of mechanism is appropriate that is not dependent on repeated interactions between the same individuals. Accordingly, the relevance of indirect reciprocity is elevated by an increase of specialization, in particular, relative to direct reciprocity. Furthermore, the relevance of indirect reciprocity increases with the rise of work in projects, as well as new opportunities in e-commerce, such as online trade. Similar to the reasoning about the increase of specialization, the decrease of repeated interactions is the point here. In projects, employees work together only for the duration of the particular project and in online trading (e.g., eBay.com and amazon.com), repeated interactions are even rarer than in project work. Accordingly, the shadow of the future is small, and hence direct reciprocity is limited. Indirect reciprocity, however, can lead to cooperation in this kind of situations.

Reputation or social status of an individual determined by its behavior reduces the problem of free-riding in indirect reciprocity. If, for example, individuals only cooperate with those of a reputation signaling that they have cooperated before, free-riding is inhibited. Thus, reputation is a key element in indirect reciprocity. The need for reputation in indirect reciprocity opens two issues. First, it demands high cognitive abilities. Second, different possibilities for determining the reputation come into play.

The high demand for cognitive abilities can be explained by examining the environmental conditions of indirect reciprocity. An audience, which is interested in the activities of others and which can understand them, is necessary for indirect reciprocity (Alexander, 1987). To this audience, or to parts of it, the source for reciprocation of the cooperative act is related. Moreover, communicating of the observed can extend indirect reciprocity's efficacy, but demands sophisticated communication abilities.

Furthermore, communication in general does not always deliver correct information, but noise is inevitable. Moreover, noise raises a possibility for opportunism. The deliberate manipulation of the reputation of others during the communication process could improve the communicator's situation. Such opportunism by influencing information of reputation has not yet been examined, although it represents a serious vulnerability for indirect reciprocity.

Overall, the issue of the cognitive abilities in the context of indirect reciprocity seems to be fulfilled by humans. Individuals can process the information observed by them or received from others. The usage of language to communicate the activities of others makes this possible. It increases the sphere of indirect reciprocity. Nowak and Sigmund (2007: 13) even say that '[i]t is likely that indirect reciprocity has provided a very selective scenario that led to cerebral expansion in human evolution' (for similar statements, see Fehr and Fischbacher, 2004: 189; Nowak and Sigmund, 2005: 1,293).

Reputation determination is the second issue of reputation under indirect reciprocity, besides the high cognitive abilities of the participants. Here, a large set of possibilities must be considered. Even by reducing reputation to a binary value, such as *good* and *bad*, multiple ways are possible to determine it. For example, the question rises whether and how the reputation of a recipient should be considered by evaluating an individual that is asked to support (i.e., to cooperate with) that recipient. Furthermore, the reputation of the donor also can influence the evaluation. To solve that problem, a specific mechanism of reputation determination, usually called *assessment rule*, is necessary.

Assessment rules can be divided into first-, second-, and third-order rules (Uchida and Sigmund, 2010). A first-order rule takes only the action of the individual into account that is asked for support. If that individual (the donor) is asked to support a recipient, the donor either supports or not. The only useful first-order assessment rule is to assign a good reputation if and only if the donor helps. This assessment rule is called *image scoring*. Nowak and Sigmund (1998a, 1998b) show that for the case of image scoring as definition for reputation determination, the condition $c/b < q$ is the prerequisite for evolutionary stability of cooperation under indirect reciprocity; where q is the probability of knowing someone's reputation. In second-order assessment, there are 16 possibilities. Here, the reputation of the recipient is considered in addition to the donor's decision. This also is either good or bad. A third-order assessment rule extends the set of second-order rules by considering the reputation of the donor. Thus, 256 possibilities exist.

Particularly in business interactions, examples for indirect reciprocity can be identified. For instance, the existence of reputation scores of the traders in online auctions, such as eBay.com, shows that indirect reciprocity takes place here. With the reputation scores, the business partners can evaluate each other and the result is a reputation as percentage value. Thus, users that have never interacted before can engage in a trade relation.

Moreover, the phenomenon of knowledge transfer should receive a special attention in the context of indirect reciprocity (Zaggl and Troitzsch, 2012). The exchange of knowledge between users in internet fora provides an example. Here, questions and answers to different topics are exchanged. A person does not get help by the same participants she or he helped before, but others in the forum help her or him. In many of these fora it is counted how many answers each user gives. Evaluation of the answers' quality is also common. Indirect reciprocity in the context of knowledge exchange is in particular important because it is almost impossible to measure the value of knowledge or information. According to the Arrow information paradox, the value of information cannot be estimated for a potential purchaser without knowing the information, but when the potential purchaser knows the information, there is no reason anymore to pay for it or give a truthful evaluation of the willingness to pay (see Arrow, 1962: 615, 1971: 152). Therefore, market mechanisms are not suitable for exchanging information and knowledge. Indirect reciprocity, however, allows exchanging information and knowledge without the need of the evaluation of its value. The low costs of sharing knowledge combined with the high benefit for the supported drives the potential of indirect reciprocity particularly for the exchange of knowledge.

Indirect reciprocity is represented in the co-citation analysis of the evolution of cooperation (Cluster IV). Furthermore, some of the reviews deal with it (Bergmüller *et al.*, 2007; Nowak, 2006; Taylor and Nowak, 2009; West *et al.*, 2006). In Cluster IV, the main publication is the *Journal of Theoretical Biology* with more than 53% of all publications. Average publication year is 2002/2003. Thus, research of indirect reciprocity can be seen as driven from the discipline of biology and to be quite young.

Strong reciprocity

Strong reciprocity is often denoted as *altruistic punishment* (e.g., Boyd *et al.*, 2003; Fehr and Gächter, 2002; Gintis *et al.*, 2008; Rockenbach and Milinski, 2006). This term is appropriate in the negative form of strong reciprocity. An individual is a strong reciprocator if it spends own resources to punish others who show adverse behavior (e.g., Fehr *et al.*, 2002). However, also individuals spending own resources to reward others who show the desired behavior are strong reciprocators.

From a game theoretic perspective, strong reciprocity works by changing the players' payoffs. If disliked behavior is punished, then the temptation for this behavior is reduced because disutility through punishment must be considered in addition to the actual payoff. The same applies reversely for strong positive reciprocity.

One argument for the importance of strong reciprocity derives from the weakness of (in)direct reciprocity; that is the dependence on future interactions. In situations, in which individuals are faced with serious problems in form of crises like famines, natural catastrophes, or wars – just when cooperation is needed the most – direct reciprocity fails because the probability of future interactions with the same individuals decreases tremendously (see also Gintis, 2000). Indirect reciprocity also depends on future interactions, not with the same individual, but with individuals from the same society. However, the continuity of that society in a crisis is questionable. Hence, indirect reciprocity neither is qualified to sustain cooperation in crises. Furthermore, it can be assumed that the costs for cooperation in such situations increase. Consequently, the general rules for the conditions of direct and indirect reciprocity predict a vanishing of cooperation in tough situations. Strong reciprocity, in contrast to (in)direct reciprocity, has the potential of sustaining cooperation even in crises.

The general problem of strong reciprocity, however, is the costs for the strong reciprocator. Although it is better for a group, and even for its individual members, when individuals who punish on own costs are present, nobody wants to pay the costs for punishing. Dreber *et al.* (2008) support this with a human subject experiment: People who tend to punish on own costs are worse off than people who forgo costly punishment. Nonetheless, strong reciprocity can be observed empirically in many different situations.

In situations, in which the costs for the strong reciprocators are low, strong reciprocity becomes more likely to be a possible outcome of evolution. This is the case in the context of hierarchies (see also Hooper *et al.*, 2010; Seymour *et al.*, 2007). When a highly ranked individual punishes lower ranked individuals, it can be assumed that the costs for the punisher are low and the impact of punishment is high. It can be speculated that costly punishment has developed under such conditions. Another possibility in addition to hierarchies to explain strong reciprocity as an evolutionary outcome is relating strong reciprocity to unpredictability. This holds in particular for the negative form (i.e., costly punishment). This is reasoned as follows. Perfectly predictable punishment is easy to outwit. One who knows how to provoke his opponent's punishment activity does consequently also know how to avoid provocation. Accordingly, she or he can maximize her or his opportunism under consideration of the punishment triggers. This makes the punishment activities inefficient. If, however, the punishing activities are hard to predict,[4] the opponent of a strong reciprocator cannot maximize his or her payoff by opportunism.

An example for strong reciprocity in the business context provides any case of a boycott by customers. Often customers avoid consuming products from companies that disappointed them in the past. If such a costumer cannot satisfy her or his demand with a substitute, her or his payoff is negatively affected.

Strong reciprocity is represented in the co-citation analysis of evolutionary game theory (Cluster 6). Furthermore, it is a topic in two of the review articles (Stevens *et al.*, 2005; West *et al.*, 2006). In Cluster 6, *Nature* is the main publication outlet. Most of the authors of the cluster's publications are economists (e.g., Samuel Bowles, Ernst Fehr, and Herbert Gintis). This makes sense against the background of the fact that strong reciprocity is a mechanism strongly tied to humans.

Overall, the implementation of strong reciprocity to foster cooperation seems to be simple. Thus, it is one of the most important cooperation mechanisms for institutional design. The effects of hierarchy and unpredictability shall be considered. Both are crucial for reducing the costs of the strong reciprocators.

Group selection

The theory of group selection, trait-group selection, or between-group selection is about how the formations of groups, such as herds, nations, or entire species, influence selection (Bergstrom, 2002; Choi and Bowles, 2007).[5] Individuals contributing to the benefit of the group are disadvantaged relative to their non-contributing peers. However, groups with a large number of contributing members are favored over groups mainly consisting of free-riders. This trade-off has received awareness early on (Darwin, 1871: 163, 166) and has been solved sometimes extremely by assigning groups an equal role as individuals in evolutionary processes (Wynne-Edwards, 1962). The formalization of group selection has driven deep insights and it has been shown that it induces cooperation under evolutionary conditions (Maynard Smith, 1964; Wade, 1978; Wilson, 1975). Traulsen and Nowak (2006) show analytically the conditions for the evolution of cooperation based on group selection working under compliance of individual selection. For that purpose, they modeled a population subdivided into groups. Only the unconditional strategies cooperation and defection characterize the individuals. Cooperators support the group members while defectors do not pay the costs for cooperation, but receive the benefits

4 Note that a minimum degree of predictability is necessary. Otherwise strong reciprocity is unconditional and loses its effectiveness entirely.

5 Group selection, in the wider sense, is represented in the co-citation analysis of the evolution of cooperation by Cluster V. Most of the review articles address it as well (Dugatkin, 1997b; Nowak, 2006; Taylor and Nowak, 2009; West *et al.*, 2006).

provided by the cooperators in the same way as the cooperators do. The individuals reproduce proportional to their payoff and the offspring is added to their parents' group. If a group has reached a certain size, it either splits up into two groups or one member is removed. When the group splits up, another group of the population is removed. Generally, cooperation evolves if the ratio of the benefit (b) generated by cooperation and its costs (c) is larger than the maximum group size (n) divided by the number of groups (m) plus one: $b/c > 1 + n/m$. An earlier model supporting group selection is provided by Wilson (1975). Furthermore, each haystack model supports group selection.

Zoology provides some empirical insights about group selection. Predator inspection in fish species helps the group to survive to the disadvantage of the inspecting individuals. It is interpreted by Dugatkin and Mesterton-Gibbons (1996) as group selection. A single fish or a small subgroup departs the fish school to approach a potential predator. The advantage for the whole group is to find out whether the predator is a threat. The disadvantages, such as reduced foraging and, even more important, the risk of death by predation, are born only by the inspecting fish. Further examples can be found by looking at warning calls in birds (e.g., Taylor *et al.*, 1990) and even in microbiology (e.g., Johnson, 1986). In humans in general, solidarity in groups, for example nations, can be considered an instance of group selection. In wars, characteristics of group selection become salient; also, conflicts between different religions can be seen as examples for group selection in humans (Choi and Bowles, 2007; van den Bergh and Gowdy, 2009). It is intuitively correct to say that whenever the threats from outside a group are rising, then the group cohesion increases. This may be the reason why groups often stress the potential of threats from outside.

The greatest potential of the mechanism of group selection is not necessarily its capability to induce cooperation directly (as it has been shown by the models of, e.g., Choi and Bowles, 2007; Maynard Smith, 1964; Traulsen and Nowak, 2006). Instead one can focus on the fertile ground groups provide for other mechanisms, such as strong and indirect reciprocity. For instance, the problem of higher order free-riding in costly punishment can vanish in the group context. Participants able to choose between groups, in which free-riders are either punished or not, prefer the punishing environment (Gürerk *et al.*, 2006). In the presence of groups with low performance, punishers are likely to be better off than the majority. For indirect reciprocity, groups provide a useful environment that ensures the conformity of assessment rules. Even more fundamentally, groups have the potential for making sure that reputation is accepted as a kind of currency for exchange. Strong and indirect reciprocity do not only profit from group selection. Reversely, the mechanisms of reciprocity push the effect of group selection by strengthening the members of the groups employing them. Also the mechanism of optional participation seems to have a strong relation with group selection. It may play a central role for the initial formation of groups. The interactions of other mechanisms with group selection make it hard to isolate it in empirical observations. However, one can speculate that the interactions shift the potential for cooperation significantly.

Costly signaling

Costly signaling is rooted in evolutionary biology as well as in economics. The development of the theory is triggered by the observation that in some animal species, especially birds, the males impose a burden on themselves by showing a salient appearance (e.g., the pheasant and the peacock). This appearance is accompanied by disadvantages, e.g., the predation risk increases because of lower escaping chances.

Zahavi (1975) first approached the question of how these animals developed their handicaps under the conditions of natural selection. Surely, sexual selection by females is an explanation, but this fact alone is not sufficient. Since sexual selection itself is subject of natural selection, an additional element is needed that explains stability. The answer is that the males of the mentioned bird species are able to demonstrate an extraordinary fitness through the handicap. Because the females cannot verify their fitness directly, the males take the burden of a handicap. The fitter males' costs for affording the handicap are relatively lower than the costs for less fit males. Consequently, the females' quality of information for evaluating potential mates is improved. Important is that the signal can only be credible if it is costly (Maynard Smith, 1991). Thus, the costs must directly be connected to the signal (Grafen, 1990).

Costly signaling plays an important role in humans as well. Besides costly signals in sexual selection, for instance luxury cars or other status symbols, examples in economics are apparent. Thus, it is possible that costly signaling has the potential to explain the Veblen (1899) effect. In business interactions as in the case of mating, information asymmetries are prevailing. Therefore, investing in a signal often makes sense for the party with hidden qualities.[6]

Spence (1973), even earlier than Zahavi (1975), used costly signaling theory in economics, more precisely in the context of job markets. The reasoning is the same. A signal must affect productivity negatively to distinguish different job applicants; in other words, it must be costly. Thus, education is an example for a signal in this context. The better-qualified applicant shows that she or he has to pay lower costs for the education. Other examples in economics are all kinds of warranties voluntarily provided by suppliers. These are costly signals send to the consumer. The supplier has more information about the product and attempts to convince the customer of the hidden quality of the product. The supplier with the better product has lower costs for offering the warranties in contrast to her or his competitors.

6 Besides that, it could be interesting to approach the idea that it may be a costly signal to abandon a costly signal. Under special circumstances the waiver of status symbols could also represent a costly signal.

Costly signaling is represented in the co-citation analysis of evolutionary game theory (Cluster 3). Furthermore, it is represented in combination with religions in the co-citation analysis of the evolution of cooperation (Cluster XI). According to Cluster 3, costly signaling is closely associated with the discipline of biology. *Journal of Theoretical Biology* and *Animal Behavior* together have a share of 75% of the publications. The economics literature mentioned is not part of the cluster.

Overall, costly signaling is highly relevant whenever unequally informed parties interact. The problem of hidden characteristics and hidden information can be overcome, at least partly, with costly signaling.

Graph selection

The idea of graph selection theory is to consider spatial structures between individuals or players. This is implicitly done in all models dealing with cooperation, but in most cases either pairwise interaction or playing the field (i.e., random pairing or global interaction with all others at the same time) is assumed. Graph selection, however, approaches to explain cooperation explicitly dependent on the spatial structure (Durrett and Levin, 1994; Ebel and Bornholdt, 2002).

In general, the modeling of spatial structures can be represented by grids (e.g., cellular automata; Hassell *et al.*, 1994; Hauert and Doebeli, 2004; Nakamaru *et al.*, 1997, 1998) or networks (Lieberman *et al.*, 2005; Santos and Pacheco, 2005; Santos *et al.*, 2006a, 2006b; Skyrms and Pemantle, 2000). The spatial structure is crucial for the origin of cooperation (Nowak and May, 1992, 1993; Nowak *et al.*, 1994). For example, Abramson and Kuperman (2001) show that even with players that are only able to play two pure unconditional strategies (cooperation and defection) and imitate the most successful neighbor's strategy, a cooperating community can emerge. Thus, cooperation can derive merely from the spatial structure. A general approximation for a rule for the stability of cooperation on regular graphs can be derived analytically (Ohtsuki and Nowak, 2006; Ohtsuki *et al.*, 2006): $b/c > k$. The number of neighbors (k) must be smaller than the benefit–cost ratio.

An early approach for spatial modeling of cooperation games comes from Axelrod (1984: 158–168). Straightforward from the exploration of optimal strategies in the prisoner's dilemma, he applied a two-dimensional structure to model territoriality. Overall, the theory of graph selection is very young and seems to develop further. Providing examples for graph selection lacks of empirical studies that are specifically designed for this. In the era of Web 2.0 and especially in online games, powerful possibilities have arisen for gathering data containing information about spatial influences on cooperation.

Graph selection is represented in both co-citation analyses (Subcluster 1.2 and Cluster I). Furthermore, only two review articles address it (Nowak, 2006; Taylor and Nowak, 2009). The mechanism plays the most prominent role in both co-citation analyses. Cluster 1.2 reveals that the discipline of physics has a crucial impact on graph selection theory's development. More than 48% of the publications are published in *Physical Review E* or in *Physical Review Letters*. In Cluster I, *Physical Review E* accounts for more than 36%. Moreover, the publications are very young. The average publication year in both clusters is between 2005 and 2006.

Set selection

Set selection, or evolutionary set theory, is closely tied to graph selection.[7] Both share the characteristic of spatiality in cooperation. The underlying representations differ, however. In set selection, the probability of interaction between two players is not homogeneously deviated over the population as it is in well-mixed models. Instead, players interact depending on their memberships in sets. Basically, a member of a set interacts with all other members of the same set. Sets can overlap, and thus, players can interact multiple times.

Evolutionary set theory is the youngest of the mechanisms explaining cooperation. Tarnita *et al.* (2009) originally introduce it with a model that includes evolutionary updating of the player's strategies and a static environment of sets. As long as the number of memberships in sets for all players is equal, an analytical approach is possible. Weakening this assumption requires simulations.

In contrast to most of the graph models, fixed spatial structures are not assumed in set selection. Instead, the set membership of the players can change. Thus, set selection models represent systems with a large degree of dynamics. This trait makes set selection a promising candidate for modeling issues of cultural evolution.

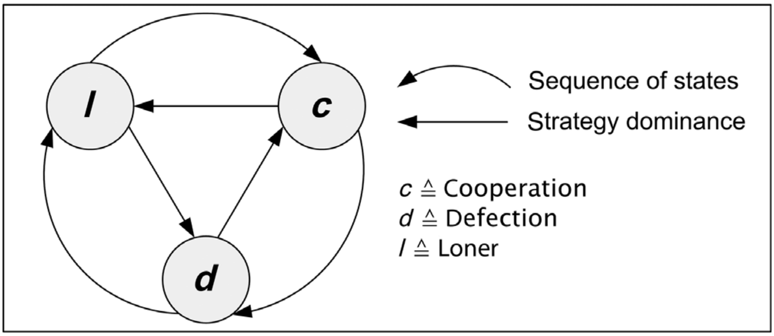
Optional participation

Optional (or voluntarily) participation is an interesting mechanism for the evolution of cooperation. Making participation voluntary can drive cooperation in dilemma games. Thereby non-participation provides a small payoff. Consequently, three basic strategies are possible: participating and contributing (cooperation); participating and not contributing (defection) i.e., free-riding; and not participating (loner). Hauert *et al.* (2002b) and Semmann *et al.* (2003) show analytically and experimentally that optional participation leads to a cyclic change of the three strategies. In consequence, the degree of cooperation is considerably higher than without the option of not participating.

The cyclic change of the strategies is a sequence of loner; followed by cooperation in the public good game, because the group can generate higher payoffs; followed by defection in the public good game, because free-riding dominates cooperation; followed again by loner, because self-supply generates

7 Set selection has been identified in the original literature (Nathanson *et al.*, 2009; Tarnita *et al.*, 2009). Neither the co-citation analyses nor the review articles address it.

Figure 1. Sequence dynamics and strategy dominance of optional participation.



higher payoffs. The relative dominance of the three strategies is opposed by this sequence of the three states. Figure 1 shows the cyclic sequence and the relation between the strategies. This sequence is sufficient to prevent a permanent invasion by free-riders.

Looking at optional participation from an abstract view reveals a fundamental logic. Therefore, it is useful to start with a normal public good game. Defection represents the only equilibrium when considering both basic pure strategies (i.e., unconditional cooperation and unconditional defection). The solution to the dilemma is introducing a third strategy with the characteristic that it dominates defection but itself is dominated by cooperation. This strategy is self-supply (i.e., the loner strategy). Overall, a transitive system is altered to an antitransitive system.

In such dynamics, which can also be found in the rock–paper–scissors game, a clear and unilateral relation of dominance between two strategies results when the third strategy dies out. This leads to a direct implication: It is not useful to over-dominate competitors, because they are the only element that protects oneself of being dominated. Furthermore, the dynamics reveal a fundamental limitation: The mechanism of cooperation can never lead to a first-best solution. The phases of not contributing to the public good game and the loner phases are essential parts of the cyclic dynamics. Here, the potential of increasing welfare by cooperation remains unexploited.

Besides its importance for cooperation, optional participation seems to be a crucial ingredient for the development of social systems including norms. This puts optional participation, in addition to its function as a mechanism of cooperation on its own, into the role of a supporter for the development of other mechanisms, such as group selection or strong reciprocity. In particular, strong reciprocity cannot evolve in populations dominated by defectors because it includes higher-order public good problems: Single or a small group of strong reciprocators are not able to establish without further explanation. However, optional instead of obligatory participation – as shown analytically and with computer-based simulations by Hauert *et al.* (2007) – makes it much more likely that societies evolve using punishment for sustaining cooperation.

Optional participation is represented in both co-citation analyses (Subcluster 1.4 and Cluster XIV). No review article addresses it. The clusters indicate that the theory of optional participation is quite young. The average publication is from 2002 to 2003. Furthermore, the clusters are very small suggesting only a few activities in the literature.

By-product mutualism

By-product mutualism is a simple but restricted mechanism for the evolution of cooperation (see, e.g., Clutton-Brock, 2009; Dugatkin, 1997a). By-product mutualism leads to cooperation because the optimal decision, even for purely self-regarding players and without further conditions, results in the optimal outcome for the opponent(s) as well. Thus, '[a]ny individual that defects (i.e., does not cooperate) in mutualistic situations will, by definition, do worse than a cooperator [..]' (Stevens *et al.*, 2005: 500). Single individuals that benefit from aggregating their power, for instance fish or bird swarms, are cases of by-product mutualism.[8]

In consequence, it is important that situations allowing for by-product mutualism are not characterized by rivalry between cooperators. From an economic perspective, by-product mutualism is nothing else than a positive externality of an individual's action. In the examples of swarms, the benefit of the externality comes only into being when enough individuals participate. However, in dyadic relations, by-product mutualism is also possible. In situations characterized by strong rivalry, in contrast, by-product mutualism is not possible. The mechanism is restricted to games characterized by the following payoff matrix: $\begin{matrix} cooperate \\ defect \end{matrix} \begin{pmatrix} R & S \\ T & P \end{pmatrix}$, with $R > T$ and $S > P$. A general rule for by-product mutualism is defined by $c < 0$; thus, the cost for the player must be negative to induce cooperation. Thus, in a strict sense, interactions based on by-product mutualism do not fulfill the definition of cooperation.

Examples from animals in the literature are numerous, particularly in fish species (Dugatkin, 1997a; Dugatkin and Mesterton-Gibbons, 1996). Utility from products with network externalities can be seen as examples for by-product mutualism in economics. The classical example for network externalities is the telephone. Each individual that acquires a telephone increases the utility of all other telephone users.

By-product mutualism has an enormous potential from the perspective of mechanism design. The behavior of players can be influenced by subsidizing particular outcomes. Accordingly, the player's payoff for the decision that leads to the result preferred by the designer is increased. From the player's perspective

8 By-product mutualism is found in four of the review articles (Bergmüller *et al.*, 2007; Dugatkin, 1997b; Stevens *et al.*, 2005; West *et al.*, 2006).

Table 3. Overview of the time series data

Mechanism	# Articles	Average age	Earliest
Kin selection	1,946	2,003.6	1964
Green beard	39	2,005.4	1981
Direct reciprocity	94	2,006.9	1961
Indirect reciprocity	567	2,008.5	1986
Strong reciprocity	637	2,009.4	2000
Group selection	1,629	2,002.3	1956
Costly signaling	537	2,003.0	1974
Spatial selection	219	2,004.5	1977
Optional participation	449	2,005.0	1973
By-product mutualism	49	2,006.0	1990

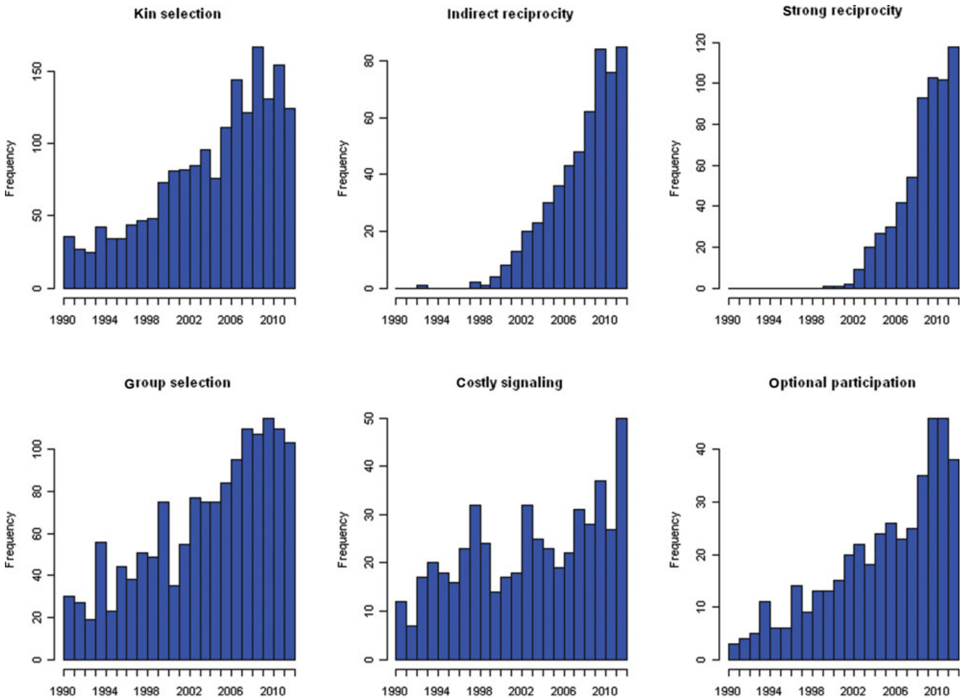
each action conform to such an incentive produces a positive by-product for the benefit of the designer. The costs for the subvention of the payoffs must not be too high to assure an efficient use. Examples for artificially created by-product mutualism are all kinds of compensation contracts to generate incentives.

Overall, by-product mutualism leads directly to a set of interesting questions in the social sciences with relevance for economics. For instance, repeated interactions between players are more realistically modeled when the payoffs vary over time. Not all interactions have the same costs and benefits. Thus, human interactions characterized by positive externalities are mixed with interactions characterized by dilemma structures. It could have a crucial effect on reciprocal behavior if the players are able to distinguish between cases of cooperation that are either costly (as in a prisoner's dilemma) or not costly for the opponent (as in the payoff matrix of the harmony game above, where $R > T$ and $S > P$). Interesting issues in this context are communication and signaling. A player, who cooperates based on by-product mutualism, could try to exaggerate his cooperativeness to convince his opponent to cooperate in the next dilemma situation. Overall, by-product mutualism plays therefore an important role, in particular, in the context of other mechanisms.

4. Bibliometric time series analysis

After collecting and reviewing the mechanisms of cooperation, a bibliometric time series analysis of them and the most important interrelations between them is conducted. For that purpose, a new data set is retrieved from the Web of Science by separately querying the mechanisms. This data set allows focusing on the individual mechanisms and their development in the literature over time. For the queries, the alternative terms (see Section 3) for the mechanisms are considered. The mechanisms of set and graph selection are merged. Table 3 gives an overview of the collected data.

Figure 2. (Colour online) Number of articles in time for the most frequently addressed mechanisms.

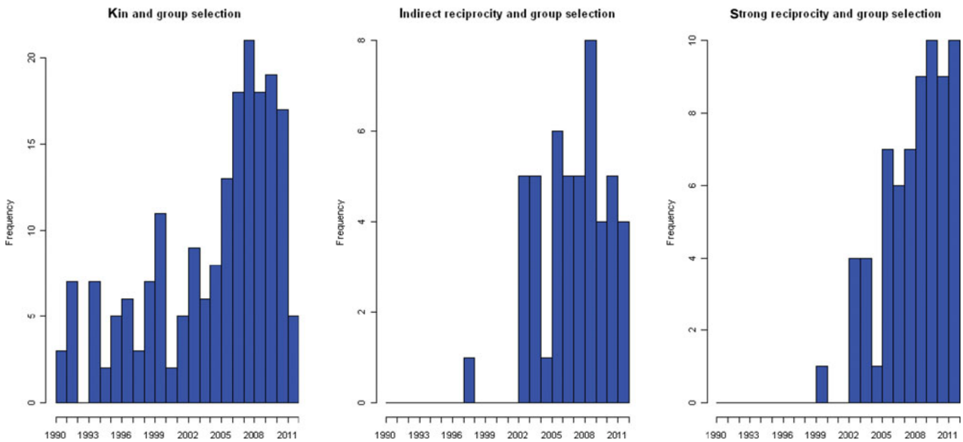


The small share of articles addressing spatial selection is remarkable, particularly in contrast to the results from the co-citation analysis where for both data sets spatiality plays the most important role. A closer look reveals that many of the articles covered in Cluster 1 and Cluster I do not use the term *selection* and consequently they are not covered by the queries. The mechanisms with the most articles are kin selection and group selection, followed by strong and indirect reciprocity, costly signaling, and optional participation. Figure 2 shows the number of articles addressing these mechanisms from 1990 to 2012.[9]

Common to all mechanisms is a strong increase in publication activity. However, kin and group selection seem to have established much earlier than the others. Costly signaling does increase only a little. Remarkable is the severe increase of the reciprocity mechanisms, particularly strong reciprocity. The earliness and the number of publications about optional participation can be seen as unexpected. The major articles formalizing optional participation are from after 2000 (Hauert *et al.*, 2002a, 2002b). However, the regard of the efficacy of voluntary participation is much older, for example, by experimental research

9 Note that direct reciprocity is underrepresented because the mechanism is often denoted as *reciprocity* without any attribute and consequently not covered by the query.

Figure 3. (Colour online) Number of articles overlapping group selection and kin selection (indirect, strong reciprocity).



(Orbell and Dawes, 1993). Group selection shows a considerable increase as well; however, it seems to stagnate at least over the past five years.

Apparently, group selection plays a central role for other mechanisms, such as indirect and strong reciprocity; and indeed, it is the mechanism with the most overlap with other mechanisms according to the data. Figure 3 shows the number of overlapping articles of group selection with the mechanisms where a significant overlap exists (these are kin selection, indirect, and strong reciprocity).

The finding of a strong overlap between group selection with indirect or strong reciprocity suits the argumentation in Section 3 very well. The interactions between these mechanisms are apparently important for cooperation. However, one intersection argued in the context of group selection is not supported by the bibliometric data of the time series analysis. Neither in the past nor currently there is a considerable amount of literature addressing group selection and optional participation. Only one publication has been identified in the process (Killingback *et al.*, 2006).

5. Framework for the mechanisms of cooperation

Up to this point, the mechanisms of cooperation have been collected and their overlaps have been examined. For the purpose of collection they are identified by using several methods complementarily. From the perspective of economics, in particular from institutional and mechanism design, the collection provides important findings. On a general level, the collection shows that the mechanisms are not homogeneously related to each other. Although some mechanisms show particular parallels, most provide fundamentally different approaches to the phenomenon of cooperation. In consequence, their potential

Figure 4. Overview of the mechanisms' categories.

Category	Spatiality	Investment	Intra-replicator Cooperation
Mechanisms	Graph selection Set selection	Direct reciprocity Indirect reciprocity Strong reciprocity Costly signaling Group selection Optional participation	Kin selection Green beard selection By-product mutualism
Description	Cooperation due to spatial structure	Investment in future benefits	Self-regarding in a strict sense (i.e., from the replicator perspective)

for application in economics and institutional design is assumed to vary substantially. To get an overview of the relationships between the mechanisms, and thus providing a framework for using the mechanisms of cooperation systematically for institutional design purposes, the collected mechanisms are framed and structured.

For that purpose, categories are developed to organize them. The motivation for the categorization is, first, to distinguish mechanisms where the stability of cooperation depends on the behavior of other individuals or not. The second criterion is to divide the mechanisms according to their dependence on temporal dynamics, or whether space is the crucial dimension. The resulting categories are *spatiality*, which contains the mechanisms where space is the crucial dimension; *investment*, which contains the mechanisms where the stability of cooperation depends on the behavior of others and time is the main dimension; and *intra-replicator cooperation*, which covers the mechanisms in which the stability of cooperation does not depend on the behavior of others. Because the mechanisms in which the stability of cooperation is independent of the behavior of others have no absolute relation to either space or time, only three categories, instead of four, are needed. An overview with a short description of the categorization of the mechanisms provides Figure 4.

The category of spatiality describes which spatial structure is prevalent. This structure determines the interaction relations of the participants for cooperation and competition. Thus, the category describes cooperation based on non-random encounter. In each cooperation model, however, a spatial structure is considered. Even in the models that do not explicitly describe their spatial structure, an implicit assumption is made because well-mixed interaction is also a kind of structure. As shown in the collection, spatiality has a crucial influence on cooperation. Furthermore, spatial or graph selection is the largest center of gravity in the field of evolutionary game theory, as well as in the field of the evolution of cooperation. The mechanisms summarized under the category of spatiality are graph selection and set selection.

In the context of institutional design, spatiality provides potential for fostering cooperation. Isolated (i.e., not in combination with mechanisms from other categories), spatiality seems to play a rather minor role by considering its practical potential to influence cooperation. However, spatiality may play a prominent role as a moderator in cooperation models.

In addition to the purely model-based papers, first empirical studies can be found in the context of spatiality (e.g., Luo *et al.*, 2011). One can expect powerful implications for practical applications from such kind of research in the future. The utilization of social networks can play a crucial role in this context. However, even with the current knowledge about the mechanisms of spatiality, some implications for the design of organizations can be derived. For instance, one could argue to recommend the minimization of the links between members of an organization to induce cooperation. This implication, however, must be put into context, not least because it may appear as counterintuitive. Certainly, several roles in an organization must be strongly connected. But it can be useful to design organizations so that small islands of cooperation can emerge.

The investment category subsumes the mechanisms based on reciprocity (i.e., direct, indirect, and strong reciprocity), costly signaling, group selection, and optional participation. They share the characteristic that individuals spend resources voluntarily to the benefit of other individuals in expectation of future rewards. Under direct reciprocity cooperation is stable if future rewards, discounted by the risk of ending the game, are higher than the current costs for cooperation. Therefore, it can be seen as an investment decision. Indirect reciprocity is also characterized by the element of current investment in future rewards. Here, reputation works as a kind of medium or currency for the investments. Strong reciprocity is also an investment in future rewards, because it is only sustainable when future rewards for the strong reciprocator or the group are realized. Costly signaling can also be subsumed under investment, because it is an investment in a signal that uncovers hidden qualities. The investment's risk may be even (part of) the signal itself. Group selection is an investment from individuals into the group's strength. The group's strength, in turn, enables benefits for the group members in the long run. In addition, optional participation should be aligned in the category of investment. Here, the risk for participating in the dilemma is the alternative for a safe benefit through self-supply. Only by taking the risk to invest, cooperation emerges. On the other hand is the loner phase, a crucial element of the mechanism. This element does not fit to the investment logic and supports an attribution to the category intra-replicator cooperation. However, overall optional participation is categorized as a mechanism of investment because the cooperation activity occurs in the cooperation phase, which underlies the logic of investment.

The mechanisms of the investment category share the characteristic of solving the dilemma and enabling cooperation by approaching the field of tension between the realization of short-term and long-term benefits. The short-term benefits can be realized by defection. This, however, excludes the chance to gain long-term benefits. The long-term benefits, in turn, are only possible by investing and forgoing short-term gains. The mechanisms of the investment category shift the weighting in this trade-off towards future and long-term gains, and accordingly towards cooperation. However, not only the field of tension between long- and short-term benefits plays a role. A requirement for cooperation is furthermore to predict and influence, at least in parts, the behavior of the opponent(s). Thus, the mechanisms of the investment category do crucially depend on the behavior of others. If an actor tries to invest, he faces always the risk of exploitation by its opponent. Because such a prediction is always accompanied by uncertainty, cooperation faces in addition the challenge of proactive defense behavior; that is, defecting before the opponent defects as a preventive strike and not for the purpose of exploitation. The higher the uncertainty about the opponent, the more likely a preventive strike occurs.

Indirect reciprocity appears to be the most interesting mechanism for the design of institutions or mechanisms. Because the investment is here not directed to a particular individual, the mechanism provides an attempt to the problem of information and knowledge exchange. Here, often unilateral relationships are apparent. Furthermore, indirect reciprocity can overcome the market failure of exchanging knowledge. Thus, indirect reciprocity provides a powerful mechanism for institutional and organizational design activities. Group selection is of similar importance. It also supports cooperation in specialized societies.

Overall, investment is the category with the most sophisticated mechanisms of cooperation. They are independent from preexisting or unalterable relations, such as kinship. Thus, it is the most important category for economics especially for institutional and mechanism design. Institutional and mechanism design efforts in the context of this category should aim for reducing the investment risk. This can be achieved by strengthening the culture of reciprocity, particularly its direct and indirect form.

The third category – the category of intra-replicator cooperation – does not rely on the principle of influencing the opponent's behavior. More specifically, it uses the criterion of the *replicator's perspective* to characterize its mechanisms. To explain that, it is useful to take the viewpoint of the replicator as the central entity of interaction (see Dawkins, 2006 [1976]). Here, the replicating entity (e.g., the parent, more precisely, the parent's genes) cooperates unilateral with its replica (e.g., the child). Because the replica is also a replicator that will soon support its own replica, the original replicator survives in new forms. Accordingly, the contained information representing the replicator only changes its 'vessel'. This shows that the replicator cooperates only with itself and basically no other individual is required. This could be seen as a contradiction of the understanding of cooperation in its meaning as the support of others on own costs. However, by switching from the perspective of the replicator back to the perspective of the individual, the definition of cooperation is fulfilled. If that switch is possible, the mechanism is among the category of intra-replicator cooperation. Thus, kin selection and green beard selection belong to the category of intra-replicator cooperation. Group selection in its strict sense (i.e., according to Wynne-Edwards) works in a similar way. If an individual supports one of its conspecifics for the purpose of the group's benefit, it is not a case of cooperation in its strict sense when the group is assumed to be the entity of selection and thus has the function of the replicator. By-product mutualism is contained in the category of intra-replicator cooperation. Here, a positive externality originated through a purely selfish act by an individual is the benefit for another individual.

For economics and institutional design, the category of intra-replicator cooperation is rather of minor importance. The required structures, such as kinship, cannot be influenced. Exceptional is by-product mutualism. As already discussed, it plays an important role for the design of incentive systems.

6. Conclusion

The mechanisms of cooperation provide a fundamental issue, which affects different fields of science, such as physics, biology, economics, and many other behavioral sciences. Using extant reviews, original literature, and co-citation analyses, a comprehensive collection of the mechanisms of cooperation is created. The collection extends the existing review articles at least by costly signaling, set selection, and optional participation. All mechanisms have been discussed and the relations between them have been characterized with a bibliometric time series analysis. Moreover, the mechanisms of cooperation have been structured into three categories. Thereby, the mechanisms in relation to each other have been examined. As a result, they cannot be seen as different pillars for cooperation. Instead their functioning is very heterogeneous. Some of them depend on characteristics of the underlying situation (e.g., by-product mutualism depends on the existence of positive externalities), others on fixed relations between the players (e.g., kin selection). The mechanisms differ in their importance from the perspective of economics and mechanism design. By structuring them into the three categories of spatiality, investment, and intra-replicator cooperation, it becomes apparent that especially the investment mechanisms are relevant for the purpose to design mechanisms in social and in particular economic systems. Furthermore, by-product mutualism should be mentioned, but except it, the category of intra-replicator cooperation plays no role for designing mechanisms. Finally, spatiality plays a role as moderator that can influence cooperation.

Although a broad spectrum of methods is employed to identify the mechanisms of cooperation, the completeness of the collection cannot be argued. According to that, it may happen that in the future some more mechanisms will become known. This is of course desirable, because each mechanism provides an additional opportunity for influencing cooperation. Furthermore, it is possible that already known mechanisms have been overlooked because of the basic limitations of the method of co-citation analysis together with the restrictions of the data sets. The additional usage of extant review articles to overcome these limitations is also restricted. Still, it can be argued that a sound approach to collect and examine the mechanisms has been designed and conducted. Furthermore, it has been suggested that the developed framework is the most appropriate one for a basic understanding of institutional design activities, but one can think of variant classifications of the mechanisms of cooperation. Thus, it cannot be argued that the presented framework is the only possible one.

Overall, the basic perspective on the mechanisms of cooperation sheds new light on the fundamentals of institutional and mechanism design. While institutional and mechanism design can be seen as the artificial creation and adaptation of mechanisms to govern people, such as employees or citizens, in social systems, the mechanisms of cooperation can be seen as natural laws framing this development.

The context of institutional and mechanism design brings the discussion to the implications. A relevant research agenda could be the detailed empirical and experimental analysis of the mechanisms of the investment category. Up to date, almost exclusive support by theoretical development of these mechanisms exists. A few experimental studies are available, but knowledge that is more detailed is needed to set up practical implications for institutional design. The huge data amount provided by multiplayer-online games can be exploited for that purpose. These games often have an economic theme. Hence, the data represent cooperation and defection decisions under stylized and even nearly laboratory-like conditions.

Furthermore, detailed implications have been identified in the review of the mechanisms. The possibility of reputation manipulations under indirect reciprocity represents a vulnerability of the mechanism's power that should be examined. Moreover, indirect reciprocity's potential for knowledge sharing and the issue of predictability of strong reciprocity should be researched.

Acknowledgements

The author wishes to thank Martin A. Nowak, David G. Rand, and two anonymous reviewers of this journal for their comments and suggestions.

References

Abramson, G. and M. Kuperman (2001), 'Social Games in a Social Network', *Physical Review E*, **63**(3): 030901.

Alexander, R. D. (1987), *The Biology of Moral Systems*, Piscataway, NJ: Aldine Transaction.
Andersson, M. (1984), 'The Evolution of Eusociality', *Annual Review of Ecology and Systematics*, **15**: 165–189.
Arrow, K. J. (1962), 'Economic Welfare and the Allocation of Resources for Invention', in NB of Economic Research (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton, NJ: Princeton University Press, pp. 609–626.
Arrow, K. J. (1971), *Essays in the Theory of Risk-Bearing*, Chicago, IL: Markham.
Aumann, R. J. (1985), 'Repeated Games', in G. R. Feiwel (ed.), *Issues in Contemporary Microeconomics and Welfare*, London: Macmillan, pp. 209–242.
Axelrod, R. (1984), *The Evolution of Cooperation*, New York: Basic Books.
Axelrod, R. and W. D. Hamilton (1981), 'The Evolution of Cooperation', *Science*, **211**(4489): 1390–1396.
Bergmüller, R., R. A. Johnstone, A. F. Russell, and R. Bshary (2007), 'Integrating Cooperative Breeding into Theoretical Concepts of Cooperation', *Behavioural Processes*, **76**(2): 61–72.
Bergstrom, T. C. (2002), 'Evolution of Social Behavior: Individual and Group Selection', *Journal of Economic Perspectives*, **16**(2): 67–88.
Bolton, G. E. and A. Ockenfels (2000), 'ERC: A Theory of Equity, Reciprocity, and Competition', *American Economic Review*, **90**(1): 166–193.
Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson (2003), 'The Evolution of Altruistic Punishment', *Proceedings of the National Academy of Sciences of the United States of America*, **100**(6): 3531–3535.
Boyd, R. and P. J. Richerson (1989), 'The Evolution of Indirect Reciprocity', *Social Networks*, **11**(3): 213–236.
Brandt, H. and K. Sigmund (2004), 'The Logic of Reprobation: Assessment and Action Rules for Indirect Reciprocation', *Journal of Theoretical Biology*, **231**(4): 475–486.
Chen, C., F. Ibekwe-Sanjuan, and J. Hou (2010), 'The Structure and Dynamics of Cocitation Clusters: A Multiple-Perspective Cocitation Analysis', *Journal of the American Society for Information Science and Technology*, **61**(7): 1386–1409.
Choi, J. K. and S. Bowles (2007), 'The Coevolution of Parochial Altruism and War', *Science*, **318**(5850): 636–640.
Clauset, A., M. E. J. Newman, and C. Moore (2004), 'Finding Community Structure in Very Large Networks', *Physical Review E*, **70**(6): 066111.
Clutton-Brock, T. (2009), 'Cooperation between Non-Kin in Animal Societies', *Nature*, **462**(7269): 51–57.
Darwin, C. (1871), *The Descent of Man and Selection in Relation to Sex*, Vol. I, London: John Murray.
Darwin, C. (1872), *The Origin of Species*, New York: Collier & Son.
Dawkins, R. (2006 [1976]), *The Selfish Gene*, Oxford: Oxford University Press.
Dreber, A., D. G. Rand, D. Fudenberg, and M. A. Nowak (2008), 'Winners Don't Punish', *Nature*, **452**(7185): 348–351.
Dufwenberg, M. and G. Kirchsteiger (2004), 'A Theory of Sequential Reciprocity', *Games and Economic Behavior*, **47**(2): 268–298.
Dugatkin, L. A. (1997a), *Cooperation among Animals: An Evolutionary Perspective*, Oxford: Oxford University Press.
Dugatkin, L. A. (1997b), 'The Evolution of Cooperation', *BioScience*, **47**(6): 355–362.
Dugatkin, L. A. and M. Mesterton-Gibbons (1996), 'Cooperation among Unrelated Individuals: Reciprocal Altruism, By-Product Mutualism and Group Selection in Fishes', *BioSystems*, **37**(1–2): 19–30.

Durrett, R. and S. Levin (1994), 'The Importance of Being Discrete (and Spatial)', *Theoretical Population Biology*, **46**(3): 363–394.
Ebel, H. and S. Bornholdt (2002), 'Coevolutionary Games on Networks', *Physical Review E*, **66**(5): 056118.
Engelmann, D. and U. Fischbacher (2009), 'Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game', *Games and Economic Behavior*, **67**(2): 399–407.
Euler, H. A. and B. Weitzel (1996), 'Discriminative Grandparental Solicitude as Reproductive Strategy', *Human Nature*, **7**(1): 39–59.
Fehr, E. and U. Fischbacher (2004), 'Social Norms and Human Cooperation', *Trends in Cognitive Sciences*, **8**(4): 185–190.
Fehr, E., U. Fischbacher, and S. Gächter (2002), 'Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms', *Human Nature*, **13**(1): 1–25.
Fehr, E. and S. Gächter (2002), 'Altruistic Punishment in Humans', *Nature*, **415**(6868): 137–140.
Fehr, E., S. Gächter, and G. Kirchsteiger (1997), 'Reciprocity as a Contract Enforcement Device: Experimental Evidence', *Econometrica*, **65**(4): 833–860.
Fehr, E., E. Kirchler, A. Weichbold, and S. Gächter (1998), 'When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets', *Journal of Labor Economics*, **16**(2): 324–351.
Fehr, E. and K. M. Schmidt (1999), 'A Theory of Fairness, Competition, and Cooperation', *Quarterly Journal of Economics*, **114**(3): 817–868.
Gintis, H. (2000), 'Strong Reciprocity and Human Sociality', *Journal of Theoretical Biology*, **206**(2): 169–179.
Gintis, H., J. Henrich, S. Bowles, R. Boyd, and E. Fehr (2008), 'Strong Reciprocity and the Roots of Human Morality', *Social Justice Research*, **21**(2): 241–253.
Gmür, M. (2003), 'Co-Citation Analysis and the Search for Invisible Colleges: A Methodological Evaluation', *Scientometrics*, **57**(1): 27–57.
Grafen, A. (1990), 'Biological Signals as Handicaps', *Journal of Theoretical Biology*, **144**(4): 517–546.
Gürerk, O., B. Irlenbusch, and B. Rockenbach (2006), 'The Competitive Advantage of Sanctioning Institutions', *Science*, **312**(5770): 108–112.
Hamilton, W. D. (1964a), 'The Genetical Evolution of Social Behaviour. I', *Journal of Theoretical Biology*, **7**(1): 1–16.
Hamilton, W. D. (1964b), 'The Genetical Evolution of Social Behaviour. II', *Journal of Theoretical Biology*, **7**(1): 17–52.
Hassell, M. P., H. N. Comins, and R. M. May (1994), 'Species Coexistence and Self-Organizing Spatial Dynamics', *Nature*, **370**(6487): 290–292.
Hauert, C. and M. Doebeli (2004), 'Spatial Structure often Inhibits the Evolution of Cooperation in the Snowdrift Game', *Nature*, **428**(6983): 643–646.
Hauert, C., de S. Monte, J. Hofbauer, and K. Sigmund (2002a), 'Replicator Dynamics for Optional Public Good Games', *Journal of Theoretical Biology*, **218**(2): 187–194.
Hauert, C., de S. Monte, J. Hofbauer, and K. Sigmund (2002b), 'Volunteering as Red Queen Mechanism for Cooperation in Public Goods Games', *Science*, **296**(5570): 1129–1132.
Hauert, C., A. Traulsen, M. A. Nowak, and K. Sigmund (2007), 'Via Freedom to Coercion: The Emergence of Costly Punishment', *Science*, **316**(5833): 1905–1907.
Hooper, P. L., H. S. Kaplan, and J. L. Boone (2010), 'A Theory of Leadership in Human Cooperative Groups', *Journal of Theoretical Biology*, **265**(4): 633–646.

Jarneving, B. (2005), 'A Comparison of Two Bibliometric Methods for Mapping of the Research Front', *Scientometrics*, **65**(2): 245–263.

Jarvis, J. U. M. (1981), 'Eusociality in a Mammal: Cooperative Breeding in Naked Mole-Rat Colonies', *Science*, **212**(4494): 571–573.

Johnson, R. (1986), 'The Group Selection of Variola Minor in the Americas: An Anthroponootic Analogue for the Myxomatosis Epizootic in Australia', *Medical Hypotheses*, **19**(4): 341–344.

Keller, L. and K. G. Ross (1998), 'Selfish Genes: A Green Beard in the Red Fire Ant', *Nature*, **394**(6693): 573–575.

Killingback, T., J. Bieri, and T. Flatt (2006), 'Evolution in Group-Structured Populations can Resolve the Tragedy of the Commons', *Proceedings of the Royal Society of London B: Biological Sciences*, **273**(1593): 1477–1481.

Kogut, B. and U. Zander (1996), 'What Firms Do? Coordination, Identity, and Learning', *Organization Science*, **7**(5): 502–518.

Laker, D. R. and M. L. Williams (2003), 'Nepotism's Effect on Employee Satisfaction and Organisational Commitment: An Empirical Study', *International Journal of Human Resources Development and Management*, **3**(3): 191–202.

Lehmann, L. and L. Keller (2006), 'The Evolution of Cooperation and Altruism: A General Framework and a Classification of Models', *Journal of Evolutionary Biology*, **19**(5): 1365–1376.

Leimar, O. and P. Hammerstein (2001), 'Evolution of Cooperation through Indirect Reciprocity', *Proceedings of the Royal Society of London B: Biological Sciences*, **268**(1468): 745–753.

Lieberman, E., C. Hauert, and M. A. Nowak (2005), 'Evolutionary Dynamics on Graphs', *Nature*, **433**(7023): 312–316.

Luo, L., N. Chakraborty, and K. Sycara (2011), 'An Evolutionary Game-Theoretic Model for Ethno-Religious Conflicts between Two Groups', *Computational and Mathematical Organization Theory*, **17**(4): 379–401.

Maynard Smith, J. (1964), 'Group Selection and Kin Selection', *Nature*, **201**(4924): 1145–1147.

Maynard Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.

Maynard Smith, J. (1991), 'Honest Signalling: The Philip Sidney Game', *Animal Behaviour*, **42**(6): 1034–1035.

Milinski, M., D. Semmann, T. C. M. Bakker, and H. J. Krambeck (2001), 'Cooperation through Indirect Reciprocity: Image Scoring or Standing Strategy?', *Proceedings of the Royal Society of London B: Biological Sciences*, **268**(1484): 2495–2501.

Milinski, M., D. Semmann, and H. J. Krambeck (2002), 'Donors to Charity Gain in both Indirect Reciprocity and Political Reputation', *Proceedings of the Royal Society of London B: Biological Sciences*, **269**(1494): 881–883.

Molander, P. (1985), 'The Optimal Level of Generosity in a Selfish, Uncertain Environment', *Journal of Conflict Resolution*, **29**(4): 611–618.

Nakamaru, M., H. Matsuda, and Y. Iwasa (1997), 'The Evolution of Cooperation in a Lattice-Structured Population', *Journal of Theoretical Biology*, **184**(1): 65–81.

Nakamaru, M., H. Nogami, and Y. Iwasa (1998), 'Score-Dependent Fertility Model for the Evolution of Cooperation in a Lattice', *Journal of Theoretical Biology*, **194**(1): 101–124.

Nathanson, C. G., C. E. Tarnita, and M. A. Nowak (2009), 'Calculating Evolutionary Dynamics in Structured Populations', *PLoS Computational Biology*, **5**(12): e1000615.

Nowak, M. A. (2006), 'Five Rules for the Evolution of Cooperation', *Science*, **314**(5805): 1560–1563.
Nowak, M. A., S. Bonhoeffer, and R. M. May (1994), 'More Spatial Games', *International Journal of Bifurcation and Chaos*, **4**(1): 33–56.
Nowak, M. A. and R. M. May (1992), 'Evolutionary Games and Spatial Chaos', *Nature*, **359**(6398): 826–829.
Nowak, M. A. and R. M. May (1993), 'The Spatial Dilemmas of Evolution', *International Journal of Bifurcation and Chaos*, **3**(1): 35–78.
Nowak, M. A., A. Sasaki, C. Taylor, and D. Fudenberg (2004), 'Emergence of Cooperation and Evolutionary Stability in Finite Populations', *Nature*, **428**(6983): 646–650.
Nowak, M. A. and K. Sigmund (1993), 'A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-For-Tat in the Prisoner's Dilemma Game', *Nature*, **364**(6432): 56–58.
Nowak, M. A. and K. Sigmund (1998a), 'The Dynamics of Indirect Reciprocity', *Journal of Theoretical Biology*, **194**(4): 561–574.
Nowak, M. A. and K. Sigmund (1998b), 'Evolution of Indirect Reciprocity by Image Scoring', *Nature*, **393**(6685): 573–577.
Nowak, M. A. and K. Sigmund (2005), 'Evolution of Indirect Reciprocity', *Nature*, **437**(7063): 1291–1298.
Nowak, M. A. and K. Sigmund (2007), 'How Populations Cohere: Five Rules for Cooperation', in R. M. May and A. R. McLean (eds.), *Theoretical Ecology: Principles and Applications*, 3rd edn, Oxford: Oxford University Press, pp. 7–16.
Nowak, M. A., C. E. Tarnita, and E. O. Wilson (2010), 'The Evolution of Eusociality', *Nature*, **466**(7310): 1057–1062.
Ockenfels, P. (1993), 'Cooperation in Prisoners' Dilemma: An Evolutionary Approach', *European Journal of Political Economy*, **9**(4): 576–579.
Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak (2006), 'A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks', *Nature*, **441**(7092): 502–505.
Ohtsuki, H. and Y. Iwasa (2004), 'How Should We Define Goodness? Reputation Dynamics in Indirect Reciprocity', *Journal of Theoretical Biology*, **231**(1): 107–120.
Ohtsuki, H., Y. Iwasa, and M. A. Nowak (2009), 'Indirect Reciprocity Provides Only a Narrow Margin of Efficiency for Costly Punishment', *Nature*, **457**(7225): 79–82.
Ohtsuki, H. and M. A. Nowak (2006), 'Evolutionary Games on Cycles', *Proceedings of the Royal Society of London B: Biological Sciences*, **273**(1598): 2249–2256.
Orbell, J. M. and R. M. Dawes (1993), 'Cooperators' Advantage, and the Option of Not Playing the Game', *American Sociological Review*, **58**(6): 787–800.
Osareh, F. (1996), 'Bibliometrics, Citation Analysis and Co-Citation Analysis: A Review of Literature I', *Libri*, **46**(3): 149–158.
Rabin, M. (1993), 'Incorporating Fairness into Game Theory and Economics', *American Economic Review*, **83**(5): 1281–1302.
Reeve, H. K., D. F. Westneat, W. A. Noon, P. W. Sherman, and C. F. Aquadro (1990), 'DNA "Fingerprinting" Reveals High Levels of Inbreeding in Colonies of the Eusocial Naked Mole-Rat', *Proceedings of the National Academy of Sciences of the United States of America*, **87**(7): 2496–2500.
Rockenbach, B. and M. Milinski (2006), 'The Efficient Interaction of Indirect Reciprocity and Costly Punishment', *Nature*, **444**(7120): 718–723.
Rubinstein, A. (1979), 'Equilibrium in Supergames with the Overtaking Criterion', *Journal of Economic Theory*, **21**(1): 1–9.

Rushton, J. P., R. J. H. Russell, and P. A. Wells (1984), 'Genetic Similarity Theory: Beyond Kin Selection', *Behavior Genetics*, **14**(3): 179–193.

Santos, F. C. and J. M. Pacheco (2005), 'Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation', *Physical Review Letters*, **95**(9): 098104.

Santos, F. C., J. M. Pacheco, and T. Lenaerts (2006a), 'Cooperation Prevails When Individuals Adjust Their Social Ties', *PLoS Computational Biology*, **2**(10): 1284–1291.

Santos, F. C., J. F. Rodrigues, and J. M. Pacheco (2006b), 'Graph Topology Plays a Determinant Role in the Evolution of Cooperation', *Proceedings of the Royal Society of London B: Biological Sciences*, **273**(1582): 51–55.

Semmann, D., H. J. Krambeck, and M. Milinski (2003), 'Volunteering Leads to Rock-Paper-Scissors Dynamics in a Public Goods Game', *Nature*, **425**(6956): 390–393.

Seymour, B., T. Singer, and R. Dolan (2007), 'The Neurobiology of Punishment', *Nature Reviews Neuroscience*, **8**(4): 300–311.

Sigmund, K. (2010), *The Calculus of Selfishness*, Princeton, NJ: Princeton University Press.

Sigmund, K. (2012), 'Moral Assessment in Indirect Reciprocity', *Journal of Theoretical Biology*, **299**(Special Issue: Evolution of Cooperation): 25–30.

Skyrms, B. and R. Pemantle (2000), 'A Dynamic Model of Social Network Formation', *Proceedings of the National Academy of Sciences of the United States of America*, **97**(16): 9340–9346.

Small, H. (1980), 'Co-Citation Context Analysis and the Structure of Paradigms', *Journal of Documentation*, **36**(3): 183–196.

Spence, M. (1973), 'Job Market Signaling', *Quarterly Journal of Economics*, **87**(3): 355–374.

Stevens, J. R., F. A. Cushman, and M. D. Hauser (2005), 'Evolving the Psychological Mechanisms for Cooperation', *Annual Review of Ecology, Evolution, and Systematics*, **36**: 499–518.

Tajfel, H. and J. C. Turner (1986), 'The Social Identity Theory of Intergroup Behavior', in S. Worchel and W. G. Austin (eds.), *Psychology of Intergroup Relations*, Chicago, IL: Nelson-Hall, pp. 7–24.

Tarnita, C. E., T. Antal, H. Ohtsuki, and M. A. Nowak (2009), 'Evolutionary Dynamics in Set Structured Populations', *Proceedings of the National Academy of Sciences of the United States of America*, **106**(21): 8601–8604.

Taylor, C. and M. A. Nowak (2009), 'How to Evolve Cooperation', in S. A. Levin (ed.), *Games, Groups, and the Global Good*, Berlin: Springer, pp. 41–56.

Taylor, R. J., D. F. Balph, and M. H. Balph (1990), 'The Evolution of Alarm Calling: A Cost-Benefit Analysis', *Animal Behaviour*, **39**(5): 860–868.

Thorne, B. L. (1997), 'Evolution of Eusociality in Termites', *Annual Review of Ecology and Systematics*, **28**: 27–54.

Traulsen, A. and M. A. Nowak (2006), 'Evolution of Cooperation by Multilevel Selection', *Proceedings of the National Academy of Sciences of the United States of America*, **103**(29): 10952–10955.

Trivers, R. L. (1971), 'The Evolution of Reciprocal Altruism', *Quarterly Review of Biology*, **46**(1): 35–57.

Uchida, S. and K. Sigmund (2010), 'The Competition of Assessment Rules for Indirect Reciprocity', *Journal of Theoretical Biology*, **263**(1): 13–19.

van den Bergh, J. C. and J. M. Gowdy (2009), 'A Group Selection Perspective on Economic Behavior, Institutions and Organizations', *Journal of Economic Behavior & Organization*, **72**(1): 1–20.

Veblen, T. (1899), *The Theory of the Leisure Class*, New York: Macmillan.

von Krogh, G., S. Haefliger, S. Spaeth, and M. W. Wallin (2012), 'Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development', *MIS Quarterly*, **36**(2): 649–676.

Wade, M. J. (1978), 'A Critical Review of the Models of Group Selection', *Quarterly Review of Biology*, **53**(2): 101–114.

West, S. A., A. S. Griffin, and A. Gardner (2006), 'Social Semantics: Altruism, Cooperation, Mutualism, Strong Reciprocity and Group Selection', *Journal of Evolutionary Biology*, **20**(2): 415–432.

Wilson, D. S. (1975), 'A Theory of Group Selection', *Proceedings of the National Academy of Sciences of the United States of America*, **72**(1): 143–146.

Wynne-Edwards, V. C. (1962), *Animal Dispersion in Relation to Social Behavior*, Edinburgh: Oliver and Boyd.

Zaggl, M. A. and K. G. Troitzsch (2012), 'Indirect Reciprocity in Social Organizations: Replication of Image Scoring and Extension to Networks', in A. Koch (ed.), *Proceedings of the 8th European Social Simulation Association Conference*, Salzburg: Universität Salzburg, pp. 291–296.

Zahavi, A. (1975), 'Mate Selection: A Selection for a Handicap', *Journal of Theoretical Biology*, **53**(1):205–214.