Social Control and the Social Contract: The Emergence of Sanctioning Systems for Collective Action

**Karl Sigmund · Christoph Hauert · Arne Traulsen ·
Hannelore De Silva**

Abstract Punishment of free-riders is generally viewed as an important factor in promoting cooperation. But since it is often costly to sanction exploiters, the emergence of such a behavior and its stability raise interesting problems. Players who do not contribute to the sanctions, but profit from the increased level of cooperation caused by them, act as "second-order exploiters" and threaten the joint enterprise. In this paper, we review the role of voluntary participation in establishing and upholding cooperation with or without punishment. In particular, we deal with two distinct forms of punishment, namely peer punishment and pool punishment, and compare their stability and their efficiency. The emergence and upkeep of collaborative undertakings can strongly depend on whether participation is voluntary or mandatory. The possibility to opt out of a joint enterprise often helps in curbing exploiters and boosting pro-social behavior.

K. Sigmund (✉)
Faculty of Mathematics, University of Vienna, 1090 Vienna, Austria
e-mail: karl.sigmund@univie.ac.at

K. Sigmund
International Institute for Applied Systems Analysis, 2361 Laxenburg, Austria

C. Hauert
Department of Mathematics, University of British Columbia, Vancouver V6T 1Z2, Canada

A. Traulsen
Max Planck Institute for Evolutionary Biology, 24306 Ploen, Germany

H. De Silva
WU (Vienna University of Economics and Business), 1090 Vienna, Austria

Introduction

A few years ago, game theorist Andrew Colman quipped that "we seem to have replaced the problem of explaining cooperation with that of explaining altruistic punishment" [20]. In this paper, we analyze mechanisms of punishment which are neither problematic nor altruistic.

A considerable amount of theoretical and experimental work, often sailing under the banner of "strong reciprocity", deals with the impact of costly punishment on the widespread human propensity to engage in collective action and contribute to joint efforts and common goods (see, e.g., [9, 15, 26, 29, 30, 32, 40, 55–58, 74, 75, 92, 107], cf. [96] for a review). Our paper surveys one particular approach to this multi-faceted issue, namely the option to turn away from the joint enterprise.

Before expanding on this topic, we would like to emphasize that the punishment of free-riders is not the only mechanism to promote cooperation. For instance, population structure (which could be provided by kinship or neighborhood) and repeated interactions assuredly play an essential role. Moreover, in many experimental and theoretical investigations, a particularly challenging scenario is assumed, namely that the collective enterprise is undertaken by a random sample of *anonymous* individuals. In many, if not most, real-life interactions, this is evidently not the case; and as soon as we allow for reputation effects and for non-random assortment in an open market for reliable partners, a large part of the problem of explaining cooperation disappears (see, e.g., [12, 55, 56, 73, 82, 83, 98]).

This being said, it seems all the more remarkable that many players engaged in experimental "public good games" are willing to pay substantial costs in order to impose fines on free-riders. The threat of punishment greatly increases the average level of pro-social contributions. But since everyone profits from an increase of the public good, those who contribute to the sanctions are providing another common good, whereas those who do not punish defectors are engaged in another form of free-riding, and may be viewed as "second-order exploiters".

From the viewpoint of evolutionary game theory, the punishment of exploiters raises interesting questions. In particular, how can such a trait emerge, and how can it be stably sustained? Evolutionary game theory is based on the assumption that successful strategies are copied preferentially, by genetic transmission, in the context of natural selection, or by social learning in cultural evolution (see, for instance, the textbooks by Weibull [105], Hofbauer and Sigmund [61], Nowak [85], Sandholm [93], and Sigmund [97]). How can strategies which are costly and provide benefits to others spread in the population?

The standard explanation offered by proponents of "strong reciprocity" is based on group selection, cf. [10, 41]. If one assumes that populations consist of separate demes, with only a limited exchange of strategies between them, then groups of pro-social punishers fare better than groups of defectors and can successfully replace them. This approach has certainly an inherent plausibility. In particular, since warring and raiding seems to take up a remarkably large place in the history and prehistory of our species (cf. [18, 44, 65, 71]), group extinction may well have been frequent enough to explain the evolution of human propensities for in-group favoritism, parochial solidarity, and even self-sacrifice for the benefit of the own tribe or nation. There is, admittedly, a vast amount of cooperation which is not related to war-like activities. But the welfare-promoting effects of such peaceful cooperation can indirectly provide an essential advantage for the competition between groups, for instance, by leading to a larger group size [69].

Is group selection a necessary mechanism for the maintenance of pro-social punishment directed against exploiters? And how can pro-social punishers get established within

a group? In this paper, we discuss a mechanism based on individual selection. It exploits the fact that in many joint enterprises, participation is voluntary rather than compulsory. It will be shown that this strongly favors the emergence of coercion. Such a result may at first almost appear as paradoxical, but it fits well with a tradition of thought pithily resumed in Hardin's maxim "mutual coercion, mutually agreed upon" [45], and which traces its roots back to social philosophers as diverse as Aristotle and Rousseau (see, e.g., [5, 6, 87, 100]). The present paper reviews and analyzes a collection of results on the effects of optional participation on contributing to joint enterprises, on the evolution of pro-social punishment and on the emergence of sanctioning institutions. This topic was first addressed in [34] and further pursued in [8, 13, 23, 28, 50, 51, 76, 94, 99, 104].

The study of the role of "policing" in promoting cooperation is by no means restricted to economics and social sciences. Indeed, it is a topic of fundamental importance in several of the "major transitions" which punctuated biological evolution, see [77]. Such transitions have led, for instance, to the cooperation of different types of molecules, to the regulation of meiotic division, to the symbiosis of proto-cells, to the emergence of multicellular organisms, and to the division of labor in social insects. It is clear that all transitions leading to higher-order units of selection are based on cooperation and thus face a basic threat from free-riders. A large number of theoretical approaches have addressed this issue, see, e.g., [33, 36, 37, 89]. But in our review, we concentrate on human collaboration. In this case, it is likely that many of the relevant types of behavior have been shaped by cultural rather than natural selection, or by gene-culture co-evolution, see e.g., [17] or [90]. The tools of evolutionary game theory are equally suited in each case: they are based on the assumption that more successful strategies are more likely to be copied. While we interpret our model in terms of cultural evolution based on social learning, we emphasize the parallels with individual-based natural selection. We show that coercion and social control of collaborative efforts can emerge spontaneously, if self-interested players can choose between engaging in a social contract or abstaining. Thus our model does not rely on a top-down approach, or require prescriptions handed down from higher authorities. On the contrary, it offers a bottom-up approach towards the emergence of sanctioning institutions.

In human societies, there exists a graduation of mechanisms of punishment, all the way from the anarchy of a laissez-faire regime up to vigilantism or strict governmental control. So far, most experiments on public goods with punishment have focused on one mechanism, namely peer punishment: *after* after the joint enterprise, individuals can impose fines on others, at a cost to themselves. The problems are obvious. In a world consisting mostly of defectors, a minority of pro-social peer-punishers bears heavy costs, and seems unlikely to be selectively favored. Conversely, in a world of cooperators, both peer-punishers and non-punishers fare equally well (since no one needs to be punished), and hence neutral drift can randomly affect their frequencies, so that the threat of punishment can dwindle to the point where defectors may invade with impunity and take over.

The very first experiment on public goods games with punishment, by Yamagishi [107], considered a different mechanism for imposing sanctions. The (anonymous) players have to decide whether or not to contribute a certain amount to a "punishment pool" *before* actually contributing to the public good game. The size of the punishment pool determines the size of the fine imposed on those who contribute least to the public good game. In [99], it is argued that this can be viewed as a step towards an institutionalized mechanism for imposing sanctions on exploiters. In particular, pool punishment is closely related to the self-financed contract enforcement games in Ostrom's "Governing the Commons" [87], and a "punishment fund" can be viewed as a rudimentary institution to uphold the common interest. Many small-scale societies use this principle, for instance, by hiring an enforcer.

When settlers pay a sheriff to keep law and order, we get an inkling of how pool punishment works. In "Governing the Commons," a series of real-life examples of self-financed contract enforcement are minutely described. They concern the provisioning and the appropriation of common resources, for instance, irrigation systems, inshore fisheries or mountain meadows, the proverbial "Commons," and discuss successful examples as well as failures.

At first glance, the emergence of pool punishment seems to run into similar obstacles as that of peer punishment, and its stability in a cooperative population is even less likely, since pool-punishers have always to contribute more than the non-punishers. However, if punishers can engage in the so-called "second-order punishment" by also punishing those who do not contribute to the sanctions, the tables are turned. We shall see that pool-punishers can establish a solid hegemony while peer-punishers cannot. Indeed, if all players cooperate, then peer-punishers cannot be distinguished from non-punishers. By contrast, pool punishers must declare themselves beforehand. Those who do not contribute to the punishment pool are just as visible as those who do not contribute to the joint enterprise. Hence, they can be punished just as well. Thus pool punishment leads more easily to an efficient second-order punishment regime, and hence to more stability.

The main aim of our paper is to show by mathematical models based on social learning that the voluntary participation in pro-social enterprises often promotes more cooperation than compulsory participation does. We analyze this for three scenarios: without punishment, with peer punishment and with pool punishment. Moreover, we distinguish whether punishment is also directed at those who do not contribute to the sanctions (second-order punishment) or not. But first, we describe the evolutionary dynamics of social learning in a general setup. Indeed, a secondary aim of this paper is to illustrate the usefulness of a simple approximation, the so-called strong imitation limit.

Evolutionary Dynamics of Social Learning

There are many ways to model social learning. The simplest is to assume that successful strategies are more likely to be imitated. We could, for instance, assume that from time to time, each player adopts the strategy of a player i chosen (within the whole population) with a probability which is an increasing function of that player's payoff P_i, for instance of the form $B + s P_i$, where B is interpreted as a "baseline fitness" (the same for all) and the parameter $s \geq 0$ measures the effect of the payoff on the probability to be imitated. Thus s measures the importance of the game: if $s = 0$, a strategy's success does not make it more or less likely to get copied. This approach, which is based on the Moran process in population genetics, was used in some of the earliest applications of social learning to evolutionary game theory, cf. [85].

We will use a different method, the so-called "pairwise comparison" approach. We assume that two players i and j are randomly chosen. Player i adopts the strategy of player j with probability $p_{i \to j}$, which is an increasing function of the payoff difference $P_j - P_i$. Similarly, with probability $p_{j \to i}$, player j adopts the strategy of player i. A popular choice for the imitation probability $p_{i \to j}$ is (see [7, 78, 101])

$$p_{i \to j} = \frac{1}{1 + \exp[s(P_i - P_j)]}, \tag{1}$$

where the "imitation strength" $s \geq 0$ measures how strictly the players are basing their decisions on payoff comparisons. In particular, for $s = 0$ or for $P_j = P_i$, a coin toss decides whether i imitates j. Small values of s correspond to a regime we call "weak imitation." In this case, imitation is basically random, but more successful players are imitated

slightly more often. For $s \to +\infty$, "strong imitation," more successful players are always imitated. Clearly, the homogeneous populations correspond to absorbing states of this stochastic process: once such a state is reached, imitation cannot produce any change.

We shall additionally assume that with a certain probability $\mu > 0$ (the so-called exploration rate), a player switches randomly to another strategy without imitating another player. The resulting Markov chain has a stationary distribution which, if the population size M is large (realistic values would be $M = 100$ or $M = 1000$), requires considerable efforts to compute numerically. In addition to individual-based computer simulations, we shall consider two limiting cases which allow an analytic treatment.

We treat the so-called "adiabatic" case which corresponds to letting the exploration rate μ tend to 0. In that case, if in a homogeneous population a single dissident arises, then its fate in the imitation process (no matter whether elimination or fixation) will be settled before the next exploratory step occurs. In this case, we only have to consider two strategies: the resident and the invader. The stationary distribution can be described by an embedded Markov chain on the pure states [38].

More precisely, let us assume that there are d strategies $1, \ldots, d$. By X_k we denote the number of players using strategy k (with $\sum X_k = M$). The homogeneous population with $X_k = M$ will be denoted by All_k. With probability $\mu/(d-1)$, a "dissident" individual switches from k to $l \neq k$. Let us denote by ρ_{kl} the probability that imitation leads to the fixation of the dissident's strategy l. This fixation probability can be computed by the formulas known from the theory of birth–death processes, see [64, 80, 103] or [85]:

$$\rho_{kl} = \frac{1}{1 + \sum_{q=1}^{M-1} \prod_{X_l=1}^{q} \frac{T_{l \to k}(X_l)}{T_{k \to l}(X_l)}}. \tag{2}$$

In our case, the probability that one out of X_l players with strategy l is chosen as a focal player and imitates one of the k-players (whose number is $M - X_l$) is given by

$$T_{l \to k}(X_l) = \frac{X_l}{M} \frac{M - X_l}{M} p_{l \to k}. \tag{3}$$

A major advantage of using the imitation probability equation (1) is that the fixation probability ρ_{kl} simplifies to

$$\rho_{kl} = \frac{1}{1 + \sum_{q=1}^{M-1} \exp\left[s \sum_{X_l=1}^{q} (P_k - P_l)\right]}, \tag{4}$$

where P_l and P_k in general depend on X_l and X_k.

In the adiabatic case, the probability of a transition from All_k to All_l is given by $\mu \rho_{kl}/(d-1)$. If the $d \times d$ transition matrix is mixing, it has a unique normalized left eigenvector to the eigenvalue 1, and this is the stationary distribution which describes the percentage of time (in the long run) spent by the population in the vicinity of the homogeneous state All_k. In [38], it is shown that the stationary distribution of the full system converges, when $\mu \to 0$, to the stationary distribution of this "embedded" Markov chain on the homogeneous states, with the transition probabilities from All_k to All_l given by $\rho_{kl}/(d-1)$ for $k \neq l$. A time scale separation argument shows that this is a good approximation if $\mu \ll M^{-2}$, as long as the game does not favor a stable coexistence of two strategies (a condition which will always be satisfied in the models considered here). For more details, we refer to [38] and [2]. We shall consider in particular the strong imitation limit, $s \to +\infty$. From (4) we see that ρ_{kl} is zero whenever $P_k > P_l$ for some value of X_k. If we always have $P_k \leq P_l$,

the fixation probability of strategy l is given by $1/m$, where m is the smallest value of X_l for which $P_k < P_l$ (and equal to M if there is no such value, so that the dynamics is only determined by stochastic drift). For a dissident who has a payoff advantage for all X_l, i.e., from its initial invasion until its fixation, we obtain $m = 1$. Such a dissident will take over the population with probability 1; if, conversely, the dissident always fares exactly as well as the resident, for all values of X_l, then the fixation probability is $1/M$.

The other limiting case we shall analyze is that of an infinitely large population size $M \to +\infty$. If we consider the strategies of types i with frequencies x_i and payoff values P_i, then $\dot{x}_i = x_i \sum_j x_j \tanh[s(P_i - P_j)/2]$ (see, e.g., [103]). In the case of weak imitation (s very small), we obtain, up to a change in velocity, the well known replicator equation (cf. [53, 54, 102] and [61])

$$\dot{x}_i = x_i \left(P_i - \sum_j x_j P_j \right). \tag{5}$$

Up to a rescaling of time, this is also obtained (for any value of s) for the previously mentioned Moran model. In the case of strong imitation (s very large), the pairwise imitation dynamics leads to the "imitate the better" dynamics [61], which is given by

$$\dot{x}_i = x_i \sum_j x_j \operatorname{sgn}[P_i - P_j] = x_i \left(\sum_{P_i > P_j} x_j - \sum_{P_i < P_j} x_j \right). \tag{6}$$

The Optional Public Good Game

An frequently used model for the "public goods game" considers groups of $N \geq 2$ anonymous individuals. Each player can decide whether or not to contribute a given amount $c > 0$ to the common pool. This amount will be multiplied by a factor $r > 1$ and then divided among all N participants, irrespective of their contribution to the public good. If N_c denotes the number of contributors, the payoff for a non-contributor is

$$rc\frac{N_c}{N}, \tag{7}$$

and that for a contributor

$$rc\frac{N_c}{N} - c. \tag{8}$$

Thus contributors receive a return rc/N from their own contribution. As long as $r < N$, this is less than c and thus raises a social dilemma: players who want to maximize their payoff should not contribute.

In this paper, we shall consider the so-called "others-only" variant (see, for instance, [107] and [23]), unless specified otherwise. In this variant, a player's contribution c is multiplied by a factor $r > 1$ and then divided among the $N - 1$ *other* players. Thus a contributor receives no return from his or her own investment. If N_c is the number of contributors, the payoff for a non-contributor is

$$rc\frac{N_c}{N-1}, \tag{9}$$

whereas a contributor obtains

$$rc\frac{N_c - 1}{N - 1} - c. \tag{10}$$

Clearly, if all players cooperate, they obtain $(r-1)c$ each. But contributors face the temptation to become non-contributors: this would improve their payoff by the amount c irrespective of the size N of the group. If all do this, however, their payoff will be 0. Hence, this game always reveals a social dilemma. This holds whenever $r > 1$, and in particular also applies for $r \geq N$.

Let us assume that the game is not compulsory (cf. [47, 48]). Some of the N players faced with the possibility of a joint enterprise may decide to abstain. We assume that they can use their time to engage in some other activity which yields an average payoff σ independently of what the other players are doing. The option to abstain from the game is of interest only if this payoff σ is larger than the payoff 0 obtained if no one contributes. Conversely, if it is larger than the payoff $(r-1)c$ obtained in a group consisting entirely of contributors, the public goods game loses all interest. Hence we shall always assume the inequality

$$0 < \sigma < (r-1)c. \tag{11}$$

It means that participating in the joint enterprise pays off if most players contribute, and fails if most defect (i.e., do not contribute). Moreover, we shall assume that there will be no public goods game if only one individual is willing to engage in the collaborative enterprise, whether as contributor or as defector. Accordingly, a solitary would-be-participant receives a payoff σ just like all the other non-participants.

In the following, we shall always assume that the population is of a given constant size M. It consists of X players who participate and contribute, Y defectors (who participate but do not contribute), and Z non-participants. By abuse of notation, we denote contributors as X-players, defectors as Y-players, etc.

Let us first consider the compulsory case, meaning that no Z-players are admitted. In a population consisting of X contributors and $M-X$ defectors, from time to time N players are sampled (without replacement). A focal X-player will have $N-1$ co-players, of whom k are contributors. Hence, this player will obtain as expected payoff

$$P_X = \sum_{k=0}^{N-1} \frac{\binom{X-1}{k}\binom{M-X}{N-1-k}}{\binom{M-1}{N-1}} \left(cr \frac{k}{N-1} - c \right). \tag{12}$$

The formula for the expected value of the hyper-geometric distribution yields the average payoff

$$P_X = cr \frac{X-1}{M-1} - c. \tag{13}$$

Similarly, defectors obtain on average

$$P_Y = \sum_{k=0}^{N-1} \frac{\binom{X}{k}\binom{M-1-X}{N-1-k}}{\binom{M-1}{N-1}} rc \frac{k}{N-1} = rc \frac{X}{M-1}. \tag{14}$$

Now let us turn to the optional case, and assume that Z players in the population are non-participants. The probability to be the only one, in a sample of N players, who is willing to participate is

$$\frac{\binom{Z}{N-1}}{\binom{M-1}{N-1}} = \frac{Z_{N-1}}{(M-1)_{N-1}}, \tag{15}$$

where we use the notation $Z_k := Z(Z-1)\cdots(Z-k+1)$.

It follows that the average payoff for defectors is

$$P_Y = \frac{Z_{N-1}}{(M-1)_{N-1}}\sigma + \left(1 - \frac{Z_{N-1}}{(M-1)_{N-1}}\right)rc\frac{M-Z-Y}{M-Z-1}, \tag{16}$$

and that for contributors

$$P_X = \frac{Z_{N-1}}{(M-1)_{N-1}}\sigma + \left(1 - \frac{Z_{N-1}}{(M-1)_{N-1}}\right)c\left(r\frac{M-Z-Y-1}{M-Z-1} - 1\right). \tag{17}$$

Clearly, we have $P_Z = \sigma$. It is easy to see that the three strategies form a Rock–Paper–Scissors cycle. Indeed, in the absence of Z-players, defectors do always better than contributors ($P_Y > P_X$). In the absence of contributors (i.e., if $X = 0$) non-participants fare better than defectors, since $P_Z \geq P_Y$ (with equality holding if and only if $Y = 1$). In the absence of defectors (i.e., if $Y = 0$), contributors fare better than non-participants since $P_Z \leq P_X$, with equality if and only if $X = 1$.

Numerical simulations show that if players use social learning to update their strategies, the three strategies X, Y and Z supersede each other in a cyclic fashion. Figure 1(a) shows a typical time-plot. As mentioned in the previous section, analytical results can be obtained for two limiting cases.

Strong imitation limit: For small exploration rates, the embedded Markov chain describing the transitions between *AllX*, *AllY* and *AllZ* is given by

$$\begin{pmatrix} 1 - \frac{1}{2}\rho_{XY} - \frac{1}{2}\rho_{XZ} & \frac{1}{2}\rho_{XY} & \frac{1}{2}\rho_{XZ} \\ \frac{1}{2}\rho_{YX} & 1 - \frac{1}{2}\rho_{YX} - \frac{1}{2}\rho_{YZ} & \frac{1}{2}\rho_{YZ} \\ \frac{1}{2}\rho_{ZX} & \frac{1}{2}\rho_{ZY} & 1 - \frac{1}{2}\rho_{ZX} - \frac{1}{2}\rho_{ZY} \end{pmatrix}. \tag{18}$$

The left eigenvector to the eigenvalue 1 is given by the transpose of

$$\begin{pmatrix} \rho_{YX}\rho_{ZX} + \rho_{YZ}\rho_{ZX} + \rho_{ZY}\rho_{YX} \\ \rho_{XY}\rho_{ZY} + \rho_{XZ}\rho_{ZY} + \rho_{ZX}\rho_{XY} \\ \rho_{XZ}\rho_{YZ} + \rho_{XY}\rho_{YZ} + \rho_{YX}\rho_{XZ} \end{pmatrix}. \tag{19}$$

Its normalization gives the stationary distribution, which can be evaluated numerically as a function of the imitation strength s, based on (4). In the limiting case of strong imitation, (4) yields $\rho_{XY} = \rho_{YZ} = 0$ and $\rho_{ZX} = 1/2$, since in a population of X- and Z-players only, there is exactly one state (namely $X = 1$) with $P_X = P_Z$. Thus the transition matrix between *AllX*, *AllY* and *AllZ* is given by

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}. \tag{20}$$

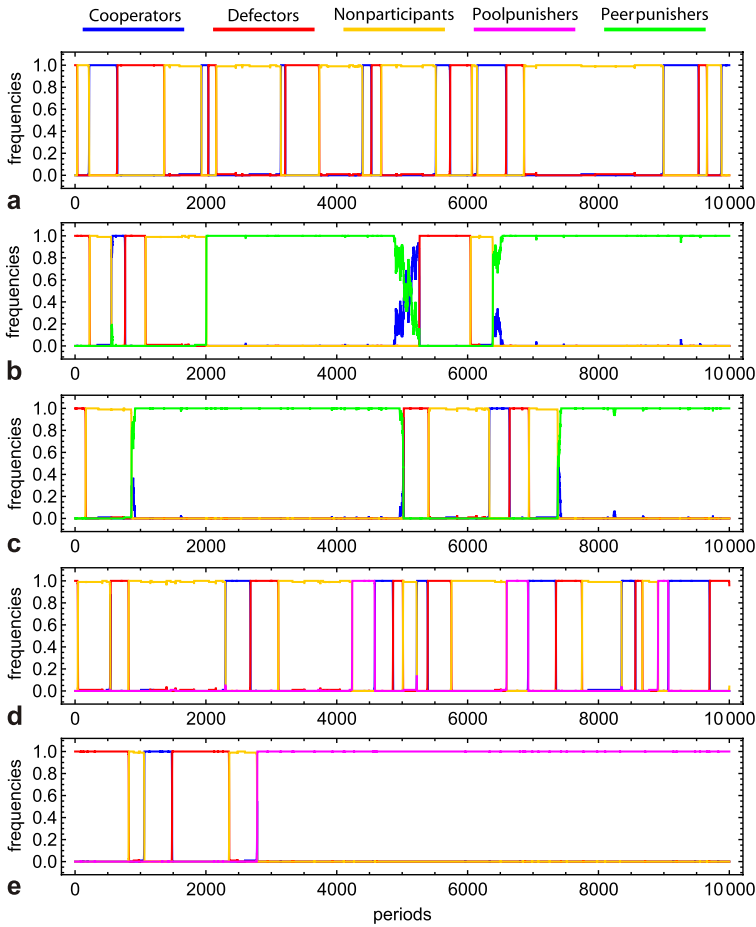The stationary distribution is given by $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$.

Infinite population limit: In this case (i.e., if $M \to \infty$), the defectors' payoff is

$$P_y = z^{N-1}\sigma + \left(1 - z^{N-1}\right)rc\frac{1-y-z}{1-z} \tag{21}$$

and that of cooperators is

$$P_x = z^{N-1}\sigma + \left(1 - z^{N-1}\right)c\left(r\frac{x}{1-z} - 1\right), \tag{22}$$

Fig. 1 Individual-based simulations of the evolution of the frequencies of strategies in a population of size M. (**a**) In the absence of punishment, contributors, defectors and non-participants supersede each other in a Rock–Paper–Scissors cycle. (**b**) If peer-punishers are introduced, they dominate for long time intervals. Sometimes their regime collapses, but is quickly re-established. In this simulation, second-order punishment was not available. (**c**) If second-order punishment is available, the dynamics is not greatly affected. (**d**) If pool-punishment is used instead of peer-punishment, the pool-punishers cannot maintain themselves in the absence of second-order punishment. (**e**) But if second-order punishment is introduced, they stably dominate the population. Parameters: $M = 100$, $N = 8$, $r = 3$, $\sigma = 0.5$, $\gamma = \beta = G = B = 0.2$, $\mu = 10^{-4}$ and $s = 100000$

with x, y, z as the relative frequencies of the three types (i.e., as the limits of X/M, etc., for $M \to \infty$). The dynamics is given by Fig. 2(b). The limiting case $s \to +\infty$, i.e., the "imitate-the-better"-dynamics (6), is specified by the rank ordering of the payoff values. If $\frac{y}{x} < \frac{c(r-1)-\sigma}{c+\sigma}$, this ordering is $P_z < P_x < P_y$. In the region $\frac{c(r-1)-\sigma}{c+\sigma} < \frac{y}{x} < \frac{rc-\sigma}{\sigma}$, the rank ordering is $P_x < P_z < P_y$, and for $\frac{rc-\sigma}{\sigma} < \frac{y}{x}$ it is $P_x < P_y < P_z$. There is no rest point in the interior of the state space S_3. The orbits are homoclinic, converging to $z = 1$ for $t \to +\infty$ and $t \to -\infty$ (see Fig. 2(a)). The case of weak selection (s very small) reduces to the replicator dynamics (5), see Fig. 2(c). In each case, there exists no fixed point in the interior of the state space, and all orbits converge to the point $z = 1$, for both $t \to +\infty$ and $t \to -\infty$.
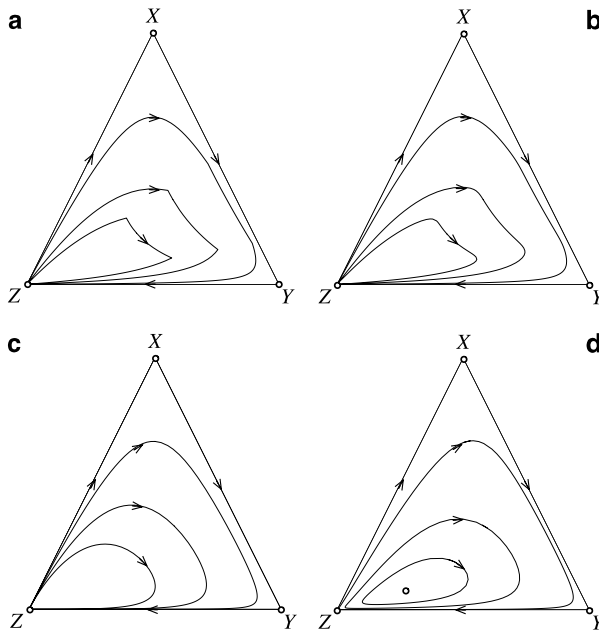
Fig. 2 Dynamics of social learning in optional public goods games in the deterministic limit of an infinitely large population. The simplex S_3 marks the phase space spanned by the three strategies cooperator, X, defector, Y, and non-participant, Z. The boundary of S_3 represents a heteroclinic cycle, reflecting the Rock–Scissors–Paper type cyclic dominance of $Z \to X \to Y \to Z$. (**a**) In the limit of strong imitation, $s \to \infty$, strategies of better performing individuals are always adopted (*imitate-the-better* dynamics). The interior of S_3 is filled with homoclinic orbits originating in the homogeneous state *AllZ*. (**b**) Dynamics generated by the microscopic imitation process of (1) for arbitrary s (here, we used $s = 20$). The noise introduced by imperfect imitation smoothes the trajectories and eliminates the non-differentiable points along the trajectories in (**a**). In the limit of weak imitation, $s \to 0$, the traditional replicator dynamics is recovered, see (**c**). A similar situation is shown in (**d**), except that we assume that part of the cooperative contributions to the public good returns to the investor, see (7) and (8). As a result a neutrally stable fixed point appears, which is surrounded by closed periodic orbits. Parameters: $N = 8$, $r = 3$, $\sigma = 0.5$

In [47, 48], the replicator dynamics is analyzed for the case that the public good game is not of the "others only" type, but with payoffs given by (7) and (8). It looks just the same for $r \le 2$ (cf. Fig. 2(c)). But if $2 < r < N$, there exists a fixed point in the interior, which is surrounded by closed orbits (cf. Fig. 2(d)). We note that in this case, if the frequency z of non-participants is large, some of the groups engaging in the public goods game will be so small that the game is no longer a social dilemma. Indeed, in that case every X-player obtains a return which is larger than the own contribution c. This is the reason why we concentrate, in the following, on the "others only" variant based on the payoffs (9) and (10). This variant is more challenging because the social dilemma always holds.

Peer Punishment

Let us now assume that W players in the population engage in peer punishment. After the public goods game, each peer-punisher imposes a fine β on each defector in his or her sample, at a cost γ. Thus if there are N_y defectors and N_w peer-punishers in the group, then each defector pays a total fine $N_w\beta$, and each punisher incurs a cost $N_y\gamma$. In addition, there

are X contributors (who participate in the public goods game, but do not punish) and Z non-participants. Since the X-players benefit from the increased contribution level caused by the W-players, they can be viewed as second-order exploiters.

If the population consists of cooperators, defectors, peer-punishers and non-participants (i.e., if $M = X + Y + Z + W$), then the average payoff for punishers is

$$P_W = P_X - \frac{(N-1)Y}{M-1}\gamma, \tag{23}$$

where P_X is given by (17). The defectors' payoff is given by the expression in (16), reduced by

$$\frac{(N-1)W}{M-1}\beta. \tag{24}$$

Computer simulations show that if (11) holds, and if the exploration rate is relatively small, then after an initial phase of oscillations of Rock–Paper–Scissors type, the population is dominated for long time-intervals by punishers. Eventually, such a W-population can be subverted by X-players (who are not punished, but do not punish either), which in turn leads rapidly to a population dominated first by defectors, and then by non-participants. The follow-up phase of cycling eventually leads to a punisher-based regime again, etc. (see Fig. 1(b)). In order to analyze this behavior, we turn to the limiting cases.

Strong imitation limit: In this case, the transition matrix for large s is

$$\begin{pmatrix} \frac{2}{3} - \frac{1}{3M} & \frac{1}{3} & 0 & \frac{1}{3M} \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{6} & 0 & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{3M} & 0 & 0 & 1 - \frac{1}{3M} \end{pmatrix}. \tag{25}$$

If the population consists only of X- and W-players, both types do equally well: this leads to the $1/M$-terms in the matrix. Conversely, X- or W-players do better than Z-players, except if $Z = M - 1$. Thus $\rho_{ZX} = \rho_{ZW} = 1/2$, and this explains the $1/6$-terms.

It is easy to see that the Markov chain (25) has a unique stationary distribution, given by

$$\frac{1}{M+8}(2, 2, 2, M+2). \tag{26}$$

This means that W-players prevail in the long run. For instance, if the population size is $M = 100$, then for almost 95% of the time, the population is dominated by peer-punishers.

If participation in the joint enterprise is compulsory, i.e., if the Z-option is not available, the transitions between the three homogeneous states $AllX$, $AllY$ and $AllW$ are given by the matrix

$$\begin{pmatrix} \frac{1}{2} - \frac{1}{2M} & \frac{1}{2} & \frac{1}{2M} \\ 0 & 1 & 0 \\ \frac{1}{2M} & 0 & 1 - \frac{1}{2M} \end{pmatrix} \tag{27}$$

and the stationary distribution is $(0, 1, 0)$. Free-riders take over.

Infinite population limit: In the limiting case of an infinitely large population, the previous payoff expressions yield

$$P_x = z^{N-1}\sigma + \left(1 - z^{N-1}\right)\left(rc\frac{x+w}{1-z} - c\right), \tag{28}$$

$$P_y = z^{N-1}\sigma + \left(1 - z^{N-1}\right)rc\frac{x+w}{1-z} - (N-1)w\beta, \qquad (29)$$

and

$$P_w = P_x - (N-1)y\gamma \qquad (30)$$

with P_x given by (28).

There is no rest point for the dynamics in the interior of the state space (i.e., for $x, y, z, w > 0$), since $P_w < P_x$ for $y > 0$. Hence all orbits converge to the boundary. Clearly, peer-punishers dominate non-participants, and peer-punishers and defectors form a bi-stable system if $c < (N-1)\beta$ (this condition is independent of γ). The states with $y = z = 0$ (i.e., when only peer-punishers and contributors are present) all correspond to rest points of the dynamics, and it can easily be shown that those with

$$w > \frac{c}{(N-1)\beta} \qquad (31)$$

are Nash-equilibria, see [51]. Depending on the initial condition, orbits converge either to such a cooperative equilibrium (which is stable, but not Lyapunov stable) or to the unstable state consisting only of non-participants. We note that Fowler [34] has obtained the same dynamics for a different, but related model.

Second-Order Peer Punishment

Let us now assume that peer-punishers engage in second-order punishment: thus they impose fines β on the X-players too, at a cost γ to themselves.

If $M = X + Y + Z + W$, the average payoff P_X for contributors is given by (17), reduced by the average fine

$$\frac{(N-1)W}{M-1}\beta\left(1 - \frac{(M-Y-2)_{N-2}}{(M-2)_{N-2}}\right) \qquad (32)$$

and the peer-punishers' payoff by (17), reduced by the average cost

$$\frac{(N-1)X}{M-1}\gamma\left(1 - \frac{(M-Y-2)_{N-2}}{(M-2)_{N-2}}\right) \qquad (33)$$

for meting out punishment. The term $(1 - ...)$ corresponds to having at least one defector in the sample (otherwise a punisher cannot be aware that some contributors do not punish).

Computer simulations with small exploration rates μ display a behavior which hardly differs from the case without second-order punishment (Fig. 1(c)).

Strong imitation limit: Indeed, it is easy to see that the behavior on the edges of the state space (where two of the four types are absent) is exactly as in the case without second-order punishment. Second order defectors cannot be noticed by peer-punishers if there are no defectors in the population. It follows that the adiabatic limit is also the same.

Infinite population limit: In the case of an infinitely large population, second-order peer punishment leads to $P_z = \sigma$,

$$P_x = z^{N-1}\sigma + \left(1 - z^{N-1}\right)\left(rc\frac{x+w}{1-z} - c\right) - (N-1)w\beta\left(1 - (1-y)^{N-2}\right), \qquad (34)$$

$$P_y = z^{N-1}\sigma + \left(1 - z^{N-1}\right)rc\frac{x+w}{1-z} - (N-1)w\beta, \qquad (35)$$
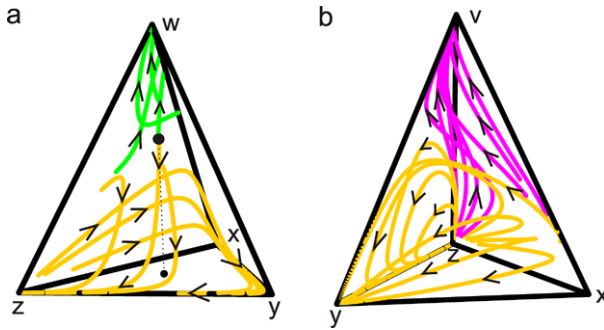
Fig. 3 Replicator dynamics for (**a**) the optional Public Goods game with peer-punishment and (**b**) the optional Public Goods game with pool punishment. In both cases, we allow for second-order punishment. Parameters for (**a**) are chosen such that an unstable fixed point exists in the interior of the state space: $N = 5$, $r = 3$, $\beta = \gamma = 0.7$, $\sigma = 1$. The states on the edge $y = z = 0$ are fixed points, of which a segment including $w = 1$ consists of Nash equilibria. For (**b**), the parameter values are $N = 8$, $r = 3$, $B = G = 0.3$, $\sigma = 0.5$

and

$$P_w = z^{N-1}\sigma + \left(1 - z^{N-1}\right)\left(rc\frac{x+w}{1-z} - c\right) - (N-1)y\gamma - (N-1)x\gamma\left(1 - (1-y)^{N-2}\right). \quad (36)$$

In contrast to the case without second-order punishment, there can exist a fixed point with all four strategies present for certain parameter values (for instance, if $\gamma = \beta = 0.7$, $c = \sigma = 1$, $r = 3$ and $N = 5$). This fixed point is unstable, however, and the orbits converge either to a mixture of W and X, or to a homogeneous Z-population (Fig. 3(a)).

Pool Punishment

Let us now assume that peer-punishers are replaced by pool-punishers. Thus we assume that V of the M players engage in pool punishment (with $M = X + Y + Z + V$). This means that when N_v of them find themselves in a sample, they participate in the public good game, and not only contribute c to the common pool, but additionally pay an extra fee G towards the punishment pool. Moreover, we assume that the fine of each exploiter will be proportional to the number of punishers, and hence of the form $N_v B$. (Since we are speaking of expected values, we may interpret $N_v B$ as the average size of the fine, and argue that if there are twice as many controls a free rider is twice as likely to get caught.)

First, we neglect the possibility of second-order punishment (i.e., punishment of non-punishers). The payoff for non-participants and for contributors is therefore unaffected, and given by σ resp. (17). The payoff for pool-punishers satisfies

$$P_V = \frac{Z_{N-1}}{(M-1)_{N-1}}\sigma + \left(1 - \frac{Z_{N-1}}{(M-1)_{N-1}}\right)\left[c\left(r\frac{M-Z-Y-1}{M-Z-1} - 1\right) - G\right]. \quad (37)$$

Indeed, punishers only pay a fee into the punishment pool if another player is willing to participate in the joint enterprise. The payoff for defectors is

$$P_Y = \frac{Z_{N-1}}{(M-1)_{N-1}}\sigma + \left(1 - \frac{Z_{N-1}}{(M-1)_{N-1}}\right)\left(cr\frac{M-Z-Y}{M-Z-1}\right) - \frac{B(N-1)V}{M-1}. \quad (38)$$

We note that the punishment term is *not* multiplied by the probability $(1 - \cdots)$ that the public goods game takes place. Indeed, if there is at least one punisher among the $N - 1$ co-players in the sample, the game will actually be played.

Instead of Condition (11), we now assume the stronger condition

$$c(r - 1) - G > \sigma, \tag{39}$$

which means that a population of pool-punishers does better than the non-participants. Computer simulations show that for a large set of parameter values, the four homogeneous states supersede each other in succession (Fig. 1(e)). We now turn to the two limiting cases.

Strong imitation limit: If the exploration rate is very small, we obtain for the strong imitation case the following transition matrix between the states *AllX*, *AllY*, *AllZ* and *AllV*:

$$\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{6} & 0 & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \end{pmatrix}. \tag{40}$$

The unique stationary distribution is given by $\frac{1}{7}(2, 2, 2, 1)$. This corresponds to two Rock–Paper–Scissors cycles, one from *AllY* to *AllZ* to *AllX* and back to *AllY* again, the other (four-membered) from *AllY* to *AllZ* to *AllV* to *AllX* and back to *AllY*. (If pool-punishers cannot invade, i.e., if (39) does not hold, then the corresponding argument leads to the stationary state $\frac{1}{5}(2, 2, 1, 0)$). We note that if the game is compulsory, i.e., if we eliminate the Z-strategy, the transition matrix between *AllX*, *AllY* and *AllV* is given by

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}. \tag{41}$$

The stationary distribution is $(0, 1, 0)$, and thus defectors prevail.

Infinite population limit: In the limiting case of an infinitely large population, the values of the expected payoffs are given by $P_z = \sigma$, P_x as in (22), and

$$P_v = z^{N-1}\sigma + \left(1 - z^{N-1}\right)\left[c\left(r\frac{x + v}{1 - z} - 1\right) - G\right], \tag{42}$$

$$P_y = z^{N-1}\sigma + \left(1 - z^{N-1}\right)cr\frac{x + v}{1 - z} - B(N - 1)v. \tag{43}$$

The inequality $P_v < P_x$ implies that there exists no interior rest point. In addition to the Rock–Paper–Scissors cycle on the face $v = 0$ (no punishment), there is another cycle leading from $z = 1$ (no participation) to $v = 1$ (all peer-punishers) to $x = 1$ (all contribute, but do not punish) to $y = 1$ (all defect) and back to $z = 1$ again, provided (39) holds. (If not, the non-participants dominate punishers.) The competition between pool-punishers and defectors (i.e., with $x = z = 0$) is bistable if and only if

$$c + G < (N - 1)B. \tag{44}$$

If the game is compulsory (i.e., if $z = 0$), then defectors always win. If the game is optional, all interior states converge to $z = 1$ for both $t \to +\infty$ and $t \to -\infty$. The limiting cases of the replicator dynamics or the "imitate the better" dynamics lead to essentially the same behavior.

Second-Order Pool Punishment

In contrast to the peer punishment case, pool punishment is strongly affected if second-order punishment is available. Indeed, let us now assume that the second-order exploiters, i.e., the X-players, are also punished by the pool-punishers. Thus their payoff is given by

$$P_X = \frac{Z_{N-1}}{(M-1)_{N-1}}\sigma + \left(1 - \frac{Z_{N-1}}{(M-1)_{N-1}}\right)c\left(r\frac{M-Z-Y-1}{M-Z-1} - 1\right) - \frac{B(N-1)V}{M-1} \quad (45)$$

and, in the infinite population limit, by

$$P_x = z^{N-1}\sigma + \left(1 - z^{N-1}\right)c\left(r\frac{x+v}{1-z} - 1\right) - B(N-1)v. \quad (46)$$

The other payoff values remain unchanged.

In this case, computer simulations show that for a large set of parameter values, an initial phase exhibiting Rock–Paper–Scissors cycles is followed by a regime dominated by the pool-punishers (see Fig. 1(e)). This regime is considerably more stable than the cooperative regime established by peer-punishers. This can also be verified in the adiabatic limit.

Strong imitation limit: If pool-punishers can invade non-participants, i.e., if (39) holds, the transition matrix between the states $AllX$, $AllY$, $AllZ$ and $AllV$ is given by

$$\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{6} & 0 & \frac{2}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (47)$$

The unique stationary distribution is $(0, 0, 0, 1)$, which means that the population is dominated by punishers. It is easy to see that in the compulsory case (i.e., no Z), defectors win.

Infinite population limit: If the game is compulsory (i.e., on the face $z = 0$), the dynamics leads either to $v = 1$ or $y = 1$. This last state can be invaded if participation is optional. Since $P_x < P_y$ in the interior of the state space, all orbits converge to the face $x = 0$, and hence either to $v = 1$ or $z = 1$, depending on the initial state (see Fig. 3(b)).

The Competition of Pool- and Peer-Punishers

We finally analyze the competition of peer and pool punishment, see [99]. Numerical simulations of populations containing all five strategies X, Y, Z, V and W show that for a large set of parameter values satisfying (39), the outcome is settled by the availability of second-order punishment. If this option is not available, then pool-punishers lose and peer-punishers prevail in the long run. With second order punishment, it is just the reverse (see Fig. 4(c)–(d)). We assume that in the case of second-order punishment, pool-punishers punish peer-punishers, since these do not contribute to the punishment pool. It seems less plausible that peer-punishers will punish pool-punishers, since these always contribute to the sanctions, and we shall accordingly not assume it here. However, we stress that this assumption does not really matter. As we shall presently see, this outcome is also reflected in the limit of small exploration rates and strong imitation.
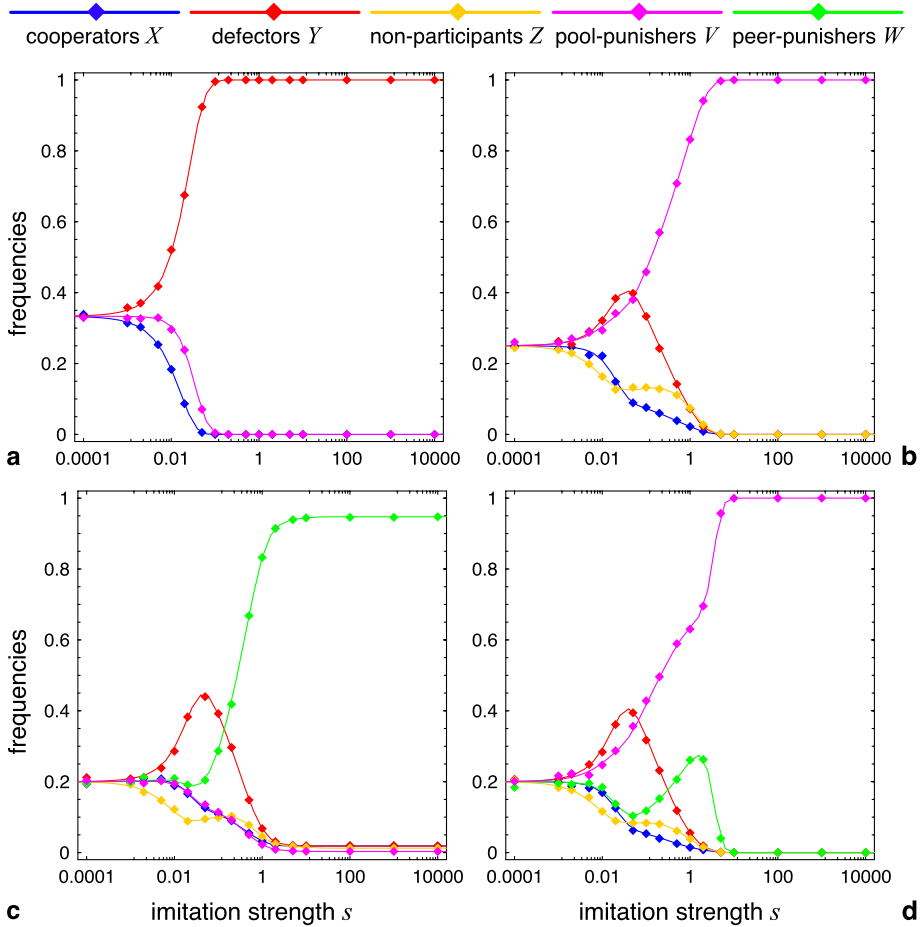
Fig. 4 Stochastic dynamics in finite population of size M of peer punishment and pool punishment in optional and compulsory public goods games with and without second order punishment, as a function of the imitation strength s. In the limit $s \to 0$, imitation is random, which results in equal average frequencies of all available strategies. For larger s, selection acts on the payoff differences between the different strategic types. The stochastic dynamics inevitably leads to absorbing homogeneous populations, where selection can no longer act. This is prevented by introducing random explorations with a small probability μ. (**a**) Compulsory public goods games with pool punishment. Because of second-order punishment both $AllY$ and $AllV$ are attracting states but for larger s, cooperation is doomed in the long run, because it is virtually impossible for a population to escape from the state of mutual defection, $AllY$. (**b**) Optional public goods games with pool punishment. Optional participation provides an escape hatch out of $AllY$ and provides recurrent opportunities to establish sanctioning institutions. For sufficiently large s, second order punishment stabilizes $AllV$ and the sanctioning institutions prevail. Without second order punishment, cooperation persists at intermediate levels, on average, as a result of the cyclic dominance of $Z \to X \to Y \to Z$ as well as of $Z \to V \to X \to Y \to Z$; but pool punishers do worst. (**c**) Competition of peer-punishers and pool-punishers in optional public goods games without second order punishment and (**d**) with second order punishment. The chosen parameters apply to situations where punishment is not very efficient. Hence defectors can dominate for noisy settings with intermediate s. For stronger imitation, peer punishment clearly dominates in the absence of second order punishment, whereas pool punishers win if second order punishment is allowed. Parameters: $M = 100$, $N = 8$, $r = 3$, $\sigma = 0.5$, $\gamma = \beta = G = B = 0.2$, $\mu = 10^{-6}$

Strong imitation limit: In this case, the transition matrix between *AllX*, *AllY*, *AllZ*, *AllV* and *AllW without* second order punishment is

$$
\begin{pmatrix}
\frac{3}{4} - \frac{1}{4M} & \frac{1}{4} & 0 & 0 & \frac{1}{4M} \\
0 & \frac{3}{4} & \frac{1}{4} & 0 & 0 \\
\frac{1}{8} & 0 & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} \\
\frac{1}{4} & 0 & 0 & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{4M} & 0 & 0 & 0 & 1 - \frac{1}{4M}
\end{pmatrix}.
\tag{48}
$$

The unique stationary distribution is $\frac{1}{3M+23}(6, 6, 4, 1, 3M+6)$. This means that the majority consists of peer-punishers. In the case *with* second order punishment, the matrix is

$$
\begin{pmatrix}
\frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 \\
0 & \frac{3}{4} & \frac{1}{4} & 0 & 0 \\
\frac{1}{8} & 0 & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} \\
0 & 0 & 0 & 1 & 0 \\
\frac{1}{4M} & 0 & 0 & 0 & 1 - \frac{1}{4M}
\end{pmatrix}.
\tag{49}
$$

The unique stationary distribution is $(0, 0, 0, 1, 0)$, which indicates that pool-punishers prevail. Again, if the game is compulsory (i.e., if there are no *Z*-players), the transition matrix between *AllX*, *AllY*, *AllV* and *AllW* is given by

$$
\begin{pmatrix}
\frac{2}{3} - \frac{1}{3M} & \frac{1}{3} & 0 & \frac{1}{3M} \\
0 & 1 & 0 & 0 \\
\frac{1}{3} & 0 & \frac{2}{3} & 0 \\
\frac{1}{3M} & 0 & 0 & 1 - \frac{1}{3M}
\end{pmatrix},
\tag{50}
$$

if there is second-order punishment. If there is no second-order punishment, the third row has to be replaced by $(\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3})$. In both cases, the unique stationary distribution is $(0, 1, 0, 0)$, so that free-riders prevail.

Discussion

Several of the "major transitions" in evolution entail the cooperation of units of selection (be they replicating molecules, genes, protocells, cells or organisms) in order to form higher units of selection. Free riders, in these different contexts, could, for instance, be outlaw genes which subvert the "fair draw" of Mendelian segregation, or cancer cells threatening the organism, or egg-laying workers in an ant colony. The suppression of free-riding is a common issue in such collaborative enterprises. This is frequently provided by mechanisms which can be viewed as punishment, or policing. To mention only a small sample of the relevant literature, we refer to Clutton-Brock and Parker [21] on punishment in animals, Ratnieks et al. [89] on policing in social insects, to Goto et al. [42] and Jander and Herre [63] on selective abortion in pollinator mechanisms, to Kiers et al. [66] on sanctions in soy beans, to Bshary and Grutter [14] on punishment of cleaner fish, to Leigh [72] on the repression of outlaw genes in meiosis, and for a general perspective, to Frank [36, 37].

We note that the model presented in Frank [35] is intimately related to the idea of pool punishment. Indeed, both approaches consider two costly traits (one for the contribution to

the common good, and the other for the repression of free-riding). In both cases, a part of the contribution is allocated towards preventing exploitation. In one case, exploitation is directly suppressed, and in the other case fines are imposed on exploiters. It is only this second case which corresponds to the common idea of punishment as a conditional mechanism which works as a threat. To explain the difference with an example: in some towns (such as Paris), automatic doors make it virtually impossible to use the metro without a valid ticket. In other towns (such as Vienna), everyone can board the underground train, but if one is caught without a ticket by a controlling officer, one has to pay a heavy fine. Frank's "suppression of competition" corresponds to the Paris system: some mechanism makes free-riding impossible, or at least very difficult. For instance, the eggs laid by rebellious ant-workers are eaten by "policing" workers. But the egg-laying workers themselves are not injured or ostracized. The "punishment pool" corresponds to the Viennese system: free-riders bear additional costs. Both ways (via automatic doors or via controlling officers) require expensive preparations. If we want to pursue the analogy, we can imagine that "peer punishment" occurs when subway passengers attack their free-riding fellow-passengers (which is also a conditional mechanism, but a rare event, to the best of our knowledge).

Human cooperation is based in a large measure on markets for reliable partners. In particular, it is often possible to opt out of a relationship that is not satisfactory. There is a considerable body of evidence that cooperation can be enhanced if participation in a joint enterprise is voluntary (see, e.g., [1, 4, 16, 43, 46–49, 52, 62, 67, 79, 86, 95]). In the present paper, we have investigated the interplay of voluntary participation and costly punishment. We have reviewed a series of papers making the (at first sight almost paradoxical) point that enforcement works less well if participation is compulsory. Most of these papers have been based on models of peer punishment: after each round, players can decide whether they inflict sanctions on the defectors. In [99], this is compared with a different model, based on so-called pool punishment: with this form of punishment, players commit themselves, even *before* the public goods interaction begins, to impose sanctions on exploiters.

A "punishment fund" can be viewed as a rudimentary institution to uphold the common interest. It is easily implemented in lab games. Actually, this was the first form of punishment used in experimental public goods games, see [107]. It seems likely that many small-scale societies use this principle, for instance, by agreeing to contribute towards hiring an enforcer. In [87] or [88], many examples of self-financed contract enforcement are described: they concern the provisioning and the appropriation of common resources, for instance, high mountain meadows (the proverbial "Commons"), irrigation systems, or inshore fisheries. For a survey of peer punishment in a small scale society, we refer to [106].

Obviously, paying for the upkeep of a police force is quite different from personally taking a hand at law-enforcement. The psychological differences are considerable. Peer punishment may be viewed as an outcome of the instinct for revenge. A player who has been exploited retaliates by harming the exploiter, even if it entails costs. It can be emotionally a most rewarding experience, and, in fact, seems intimately linked to reward-centers in the deeper part of our brain (see, e.g., [91] or [24]). If a small child bangs against a door-post, it may feel the urge to kick it. Pool punishment is more cool-headed: it is based on foresight rather than on anger, and likely to be correlated with activities in the frontal lobes of the brain, which are usually associated with rational decisions. In this sense, pool punishment is more detached than peer punishment. The commitment to punish is made *before* it is known against whom it will be directed. It is impersonal. In contrast, peer punishment seems a direct manifestation of the urge to defend one's own personal interests. It may be significant, in this context, that the first police force in ancient Athens, an early manifestation of a punishment pool, consisted of slaves, i.e., individuals who (in the eyes of their contemporaries) had no own interests to defend.

It is almost unavoidable that a sanctioning institution views non-contribution as a grave form of defection, and thus should turn to second-order punishment. It should moreover strive to suppress peer-punishment and to ensure a monopoly. We see pool punishment as a transitory mechanism, at an early stage of the evolution of a government. It seems plausible that different types of joint enterprises within a society would eventually use the same policing forces, or syndics, or law-courts, and thus converge to a form of local government. But this need not always be the case. Examples from [87] show that the local enforcement of irrigation rules in Spain or the Philippines remained, to a large extent, independent of the state authorities. Moreover, it can spectacularly break down if top–down rules are superimposed.

Compared with peer punishment, pool punishment has a serious disadvantage: it entails fixed costs even if every one in the population cooperates. Peer punishment, in that case, is cost-free. Many of the early experiments on public goods with peer punishment ended after a few rounds, and while punishment invariably raised the propensity to cooperate, it entailed costs. The overall income was therefore often less than in the absence of punishment, see, e.g., [31] and [25]. But as shown in [39], if the number of rounds is sufficiently large, cooperation becomes so common that peer punishment hardly ever occurs; and then, the overall income is higher than in the absence of punishment. Pool punishment is less efficient than peer punishment, in this regard, since the costs of pool punishment arise even if there are no exploiters to be punished (just as the upkeep of a constabulary causes costs even if no crimes occur).

On the other hand, pool punishment seems less likely to be perverted by "anti-social" or "perverse" punishment directed at the contributors, cf. [19, 59, 84]. In [27], it is shown that if players can vote, before the public good interaction, whether the subsequent peer punishment should be applied to defectors only, or could be used without restrictions, then anti-social punishment becomes very rare. This seems to indicate that when players set up a punishment pool, they would direct it only against free-riders.

Moreover, peer punishment is ill-suited to second-order punishment. It has often been noted that second-order peer punishment could be subverted by third order exploiters etc. For sufficiently high m, mth order punishment is unlikely to ever be used, and hence can be subverted by $(m - 1)$-order free riders, so that the whole mechanism ultimately unravels. Pool punishment, as shown by our models, is much more conducive to second-order punishment. The contribution to the punishment pool can be viewed as a costly signal, and hence cannot easily be faked. The claim that peer punishment does rarely extend to second-order punishment has also been supported empirically by Kiyonari et al. [68] (for a dissenting opinion, see [3]). It is sometimes asserted that the harsh punishment meted out to strike-breakers is an example of second-order punishment, but we contend that from the workers' perspective, the joint undertaking is not the firm, but the strike, and hence punishment of strike-breakers would rank as a first-order punishment. Moreover, a strike often involves trade-unions, and there (just as with guilds, and other forms of self-governing institutions), pool punishment is more likely to apply.

It is probable that many real-life defectors are opportunistic, and therefore would switch to cooperation if the threat of punishment looms large. In this paper, we have not considered opportunistic strategies, in order to make our models more easily comparable with the bulk of work on punishment on laboratory games, which usually occur under conditions of anonymity. But there is no doubt that opportunism is a very important factor. Players who see that others, in their group, are building up a punishment funds, ought to be strongly motivated to avoid punishment by contributing both to the public good and, if need be, to the punishment pool. Players who incline to contingent cooperation are encouraged to collaborate if they see others committing themselves to monitor contributions and to enforce

them. By contrast, in the peer punishment case, would-be exploiters are not able to gage the risks which they face. This is an immediate consequence of our anonymity assumption. Exploiters may be interacting with a phalanx of punishers, but will only notice this when it is too late. With peer-punishment, an opportunistic switch to cooperation can only happen if players know beforehand that there are punishers in their group. This requires that players are not anonymous, but have a reputation. Such reputation effects have been studied in [22, 49, 60, 98].

A further topic of great interest is graduated punishment, see, e.g., [87] and [106] and, for a theoretical treatment, [81] or [57]. Another highly plausible assumption, which we have not pursued here, is that punishment is contingent upon the number of punishers. Essentially, peer-punishers ought to make sure that they will share the cost with a sufficiently large set of other punishers. This has been modeled by Boyd, Gintis and Bowles [11] in the context of a repeated game. Punishers signal in a first stage their intent to punish defectors, and punish if enough others are signaling, too. If there are increasing returns to scale, and if defectors contribute in subsequent rounds after having been punished, punishment can be stably maintained.

Among other approaches to the emergence of institutions, we refer in particular to [70], where the "dilemma of endogenous institution formation" is discussed. A solution modeled on the Kyoto protocol is analyzed theoretically and experimentally. It differs from our approach in several respects. In particular, only members of the sanctioning institution can be punished. The "institution formation" relies on a complex process of quorum-sensing. In a first stage, players have to indicate their willingness to participate. In the second stage, they can ratify their intention to join, or not: and the institution is implemented if and only if all who indicated their willingness actually did ratify their intention. Players unwilling to join the punishment pool can free-ride with impunity on the (non-excludable) public good. This model describes well how modern states create sanctioning mechanisms, but it seems too sophisticated to apply to simple small-scale societies.

In the experimental game described in [43], players are given, between rounds, the opportunity to choose between a public goods game with and a public goods game without punishment. Significantly, the majority first tends to opt for the treatment without punishment, but then, after a few rounds, convert to the treatment with punishment. This experiment used peer punishment. It would be interesting to repeat it for pool punishment, and even more to allow both forms of punishment to compete: according to the model presented in this paper, the opportunity to mete out second-order punishment should decisively affect the outcome.

Our model shows that by simply imitating what succeeds, individuals can spontaneously adopt a self-governing institution to monitor contributions and sanction free riders. It needs no higher authority and no great feats of planning: trial and error and the imitation of successful examples can do the job.

References

1. Aktipis CA (2004) Know when to walk away: Contingent movement and the evolution of cooperation. J Theor Biol 231:249–260
2. Antal T, Scheuring I (2006) Fixation of strategies for an evolutionary game in finite populations. Bull Math Biol 68:1923–1944

3. Baranski B et al (2006) Higher-order punishment and the evolution of cooperation. In: GECCO'06, Seattle, Washington
4. Batali J, Kitcher P (1995) Evolution of altruism in optional and compulsory games. J Theor Biol 175:161–171
5. Binmore K (1994) Game theory and the social contract. Playing fair, vol 1. MIT Press, Cambridge
6. Binmore K (1998) Game theory and the social contract. Just playing, vol 2. MIT Press, Cambridge
7. Blume LE (1993) The statistical mechanics of strategic interaction. Games Econ Behav 5:387–424
8. Boyd R, Mathew S (2007) A narrow road to cooperation. Science 316:1858–1859
9. Boyd R, Richerson PJ (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol Sociobiol 13:171–195
10. Boyd R, Gintis H, Bowles S, Richerson P (2003) The evolution of altruistic punishment. Proc Natl Acad Sci USA 100:3531–3535
11. Boyd R, Gintis H, Bowles S (2010) Coordinated punishment of defectors sustains cooperation and can proliferate when rare. Science 328:617–620
12. Brandt H, Hauert C, Sigmund K (2003) Punishment and reputation in spatial public goods games. Proc R Soc B 270:1099–1104
13. Brandt H, Hauert C, Sigmund K (2006) Punishing and abstaining for public goods. Proc Natl Acad Sci 103:495–497
14. Bshary R, Grutter AS (2005) Punishment and partner switching cause cooperative behavior in a cleaning mutualism. Biol Lett 1:396–399
15. Cant MA, Johnstone RA (2006) Self-serving punishment and the evolution of cooperation. J Evol Biol 19:1383–1385
16. Castro L, Toro MA (2008) Iterated prisoner's dilemma in an asocial world dominated by loners, not by defectors. J Theor Biol 74:1–5
17. Cavalli-Sforza LL, Feldman M (1981) Cultural transmission and evolution. Princeton UP, Princeton
18. Choi JK, Bowles S (2007) The co-evolution of parochial altruism and war. Science 318:636–640
19. Cinyabuguma M, Page T, Putterman L (2006) Can second-order punishment deter perverse punishment? Exp Econ 9:265–279
20. Colman, A (2006) The puzzle of cooperation. Nature 440:744–745
21. Clutton-Brock TH, Parker GA (1995) Punishment in animal societies. Nature 373:209–216
22. De Silva H, Sigmund K (2009) Public good games with incentives: The Role of reputation. In: Levin SA (ed) Games, groups and the global good. Springer series in game theory. Springer, Berlin, pp 85–114
23. De Silva H, Hauert C, Traulsen A, Sigmund K (2010) Freedom, enforcement, and the social dilemma of strong altruism. J Evol Econ 20:203–217
24. deQuervain DJF, Fischbacher U, Treyer V, Schelthammer M, Schnyder U, Buck A, Fehr E (2004) The neural basis of altruistic punishment. Science 305:1254–1258
25. Dreber A, Rand DJ, Fudenberg D, Nowak MA (2008) Winners don't punish. Nature 452:348–351
26. Egas M, Riedl A (2008) The economics of altruistic punishment and the maintenance of cooperation. Proc R Soc B 275:871–878
27. Ertan A, Page T, Putterman L (2009) Who to punish? Individual decisions and majority rule in mitigating the free rider problem. Eur Econ Rev 53:495–511
28. Falk A, Fehr E, Fischbacher U (2002) Appropriating the commons: A theoretical explanation. In: Ostrom et al (eds) The drama of the commons. National Academy Press, Washington, pp 157–191
29. Fehr E, Gächter S (2000) Cooperation and punishment in public good experiments. Am Econ Rev 90:980–994
30. Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415:137–140
31. Fehr E, Rockenbach B (2003) Detrimental effects of sanctions on human altruism. Nature 422:137–140
32. Fehr E, Fischbacher U (2004) Third-party punishment and social norms. Evol Hum Behav 25:63–87
33. Foster KR (2004) Diminishing returns in social evolution: the not-so-tragic commons. J Evol Biol 17:1058–1072
34. Fowler JH (2005) Altruistic punishment and the origin of cooperation. Proc Natl Acad Sci USA 102:7047–7049
35. Frank SA (1995) Mutual policing and repression of competition in the evolution of cooperative groups. Nature 377:520–522
36. Frank SA (1996) Policing and group cohesion when resources vary. Anim Behav 52:1163–1169
37. Frank SA (2003) Repression of competition and the evolution of cooperation. Evolution 57:693–705
38. Fudenberg D, Imhof L (2005) Imitation processes with small mutations. J Econ Theory 131:251–262
39. Gächter S, Renner E, Sefton M (2008) The long-run benefits of punishment. Science 322:1510
40. Gardner A, West SA (2004) Cooperation and punishment, especially in humans. Am Nat 164:753–764
41. Gintis H, Bowles S, Boyd R, Fehr E (eds) (2005) Moral sentiments and material interests: The foundations of cooperation in economic life. MIT Press, Cambridge

42. Goto R, Okamoto T, Kiers T, Kawakita A, Kato M (2010) Selective flower abortion maintains moth cooperation in a newly discovered pollinator mutualism. Ecol Lett 13:321–329
43. Gürerk O, Irlenbush B, Rockenbach B (2006) The competitive advantage of sanctioning institutions. Science 312:108–111
44. Hamilton WD (1975) Innate social aptitudes of man: An approach from evolutionary genetics. In: Fox R (ed) Biosocial anthropology. Malaby Press, London, pp 133–153. See also Hamilton WD (1996) Narrow roads to gene land, vol I. Freeman, New York
45. Hardin G (1968) The tragedy of the commons. Science 162:1243–1248
46. Hauert C, Szabo G (2003) Prisoner's dilemma games and public goods games in different geometries: Compulsory vs voluntary interactions. Complexity 8:31–38
47. Hauert C, DeMonte S, Hofbauer J, Sigmund K (2002) Volunteering as a Red Queen mechanism for cooperation. Science 296:1129–1132
48. Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Replicator dynamics for optional public goods games. J Theor Biol 218:187–194
49. Hauert C, Haiden N, Sigmund K (2004) The dynamics of public goods. Discrete Contin Dyn Syst B 4:575–585
50. Hauert C, Traulsen A, Nowak MA, Brandt H, Sigmund K (2007) Via freedom to coercion: The emergence of costly punishment. Science 316:1905–1907
51. Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K (2008) Public goods with punishment and abstaining in finite and infinite populations. Biol Theory 3:114–122
52. Hayashi N, Yamagishi T (1998) Selective play: Choosing partners in an uncertain world. Pers Soc Psychol Rev 2:276–289
53. Helbing D (1996) A stochastic behavioral model and a 'microscopic' foundation of evolutionary game theory. Theory Decis 40:149–179
54. Helbing D (1993) Boltzmann-like and Boltzmann–Fokker–Planck equations as a foundation of behavioral models. Physica A 196:546–573
55. Helbing D, Szolnoki A, Perc M, Szabó G (2010) Evolutionary establishment of moral and double moral standards through spatial interactions. PLoS Comput Biol 6:e1000758
56. Helbing D, Szolnoki A, Perc M, Szabó G (2010) Defector-accelerated cooperativeness and punishment in public goods games with mutations. Phys Rev E 81:057104
57. Helbing D, Szolnoki A, Perc M, Szabó G (2010) Punish, but not too hard: How costly punishment spreads in the spatial public goods game. New J Phys 12:083005
58. Henrich J et al (2006) Costly punishment across human societies. Science 312:176–177
59. Herrmann B, Thoeni C, Gächter S (2008) Antisocial punishment across societies. Science 319:1362–1367
60. Hilbe C, Sigmund K (2010) Incentives and opportunism: from the carrot to the stick. Proc R Soc B 277:2427–2433
61. Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge UP, Cambridge
62. Izquierdo SS, Izquierdo LR, Vega-Redondo F, Weibull JW (2009) The option to leave: Conditional dissociation in the evolution of cooperation. Working paper
63. Jander KC, Herre EA (2010) Host sanctions and pollinator cheating in the fig tree-fig wasp mutualism. Proc R Soc B 277:1481–1488
64. Karlin S, Taylor HM (1975) A first course in stochastic processes. Academic Press, New York
65. Keeley L (1996) War before civilisation. Oxford UP, Oxford
66. Kiers ET, Rousseau RA, West SA, Denison RF (2003) Host sanctions and the legume-rhizobium mutualism. Nature 425:78–81
67. Killingback T, Bieri J, Flatt T (2006) Evolution in group-structured populations can solve the tragedy of the commons. Proc R Soc B 273:1477–1481
68. Kiyonari T, Barclay P, Wilson M, Daly D (2004) Second order punishment in one-shot prisoner's dilemma. Int J Psychol 39:329–334
69. Kokko H, Johnstone RA, Clutton-Brock TH (2001) The evolution of cooperative breeding through group augmentation. Proc R Soc Lond B 268:187–196
70. Kosfeld M, Okada A, Riedl A (2009) Institution formation in public goods games. Am Econ Rev 99:1335–1395
71. LeBlanc SA (2003) Constant battles. St Martins, New York
72. Leigh EG (1971) Adaptation and diversity. Freeman, San Francisco
73. Levin SA (ed) (2009) Games, groups, and the global good. Springer, New York
74. Lehmann L, Rousset F, Roze D, Keller L (2007) Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. Am Nat 170:21–36

75. Masclet D, Noussair C, Tucker S, Villeval M (2003) Monetary and non-monetary punishment in the voluntary contributions mechanism. Am Econ Rev 93:366–380
76. Mathew S, Boyd R (2009) When does optional participation allow the evolution of cooperation? Proc R Soc B 276:1167–1174
77. Maynard Smith J, Szathmary E (1995) The major transitions in evolution. Freeman, Oxford
78. McFadden D (1981) Structural discrete probability models derived from theories of choice. In: Manski CF, McFadden DL (eds) Structural analysis of discrete data and econometric applications. MIT Press, Cambridge
79. Miller RR (1967) The option of not playing the game. J Pers Soc Psychol 6:150–156
80. Moran PAP (1962) The statistical processes of evolutionary theory. Clarendon Press, Oxford
81. Nakamaru M, Dieckmann U (2009) Runaway selection for cooperation and strict-and-severe punishment. J Theor Biol 257:1–8
82. Nakamaru M, Iwasa Y (2005) The evolution of altruism by costly punishment in lattice-structured populations: Score-dependent viability vs score-dependent fertility. Evol Ecol Res 7:853–870
83. Nakamaru M, Iwasa Y (2006) The evolution of altruism and punishment: Role of the selfish punisher. J Theor Biol 240:475–488
84. Nikiforakis N (2008) Punishment and counter-punishment in public good games. J Polit Econ 92:91–112
85. Nowak MA (2006) Evolutionary dynamics. Harvard UP, Cambridge
86. Orbell JH, Dawes RM (1993) Social welfare, cooperator's advantage, and the option of not playing the game. Am Soc Rev 58:787–800
87. Ostrom E (1990) Governing the commons: The evolution of institutions for collective action. Cambridge UP, Cambridge
88. Ostrom E, Walker J (2003) Trust and reciprocity: Interdisciplinary lessons from experimental research. Russell Sage Foundation, Thousand Oaks
89. Ratnieks FLW, Foster K, Wenseleers T (2006) Conflict resolution in insect societies. Ann Rev Entomol 51:581–608
90. Richerson P, Boyd R (2005) Not by genes alone: How culture transformed human evolution. Chicago UP, Chicago
91. Rilling JK, Gutmann DA, Zeh TR, Pagnoni G, Berns GS, Kilts CD (2002) A neural basis for social cooperation. Neuron 35:395–405
92. Rockenbach B, Milinski M (2006) The efficient interaction of indirect reciprocity and costly punishment. Nature 444:718–723
93. Sandholm WH (2010) Population games and evolutionary dynamics. MIT Press, Cambridge
94. Sasaki T, Okada I, Unemi T (2007) Probabilistic participation in public goods games. Proc R Soc B 274:2639–2643
95. Semmann D, Krambeck HJ, Milinski M (2003) Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature 425:390–393
96. Sigmund K (2007) Punish or perish? Retaliation and collaboration among humans. Trends Ecol Evol 22:593–600
97. Sigmund K (2010) The calculus of selfishness. Princeton UP, Princeton
98. Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. Proc Natl Acad Sci USA 98:10757–10762
99. Sigmund K, De Silva H, Traulsen A, Hauert C (2010) Sanctions on the commons: Social learning and the social contract. Nature 466:861–863
100. Skyrms B (2004) The stag hunt and the evolution of social structure. Cambridge UP, Cambridge
101. Szabo G, Toke C (1998) Evolutionary prisoner's dilemma game on a square lattice. Phys Rev E 58:69–73
102. Traulsen A, Claussen J, Hauert C (2005) Coevolutionary dynamics: From finite to infinite populations. Phys Rev Lett 95:23870
103. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. Phys Rev E 74:01190
104. Traulsen A, Hauert C, De Silva H, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. Proc Natl Acad Sci 106:709–706
105. Weibull J (1995) Evolutionary game dynamics. MIT Press, Cambridge
106. Wiessner P (2005) Norm enforcement among the Ju/'hoansi Bushmen: A case of strong reciprocity? Hum Nat 16:115–145
107. Yamagishi T (1986) The provision of a sanctioning system as a public good. J Pers Soc Psychol 51:110–116