Journal of Theoretical Biology 265 (2010) 624–632

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Anti-social punishment can prevent the co-evolution of punishment and cooperation

David G. Rand [a,b,*], Joseph J. Armao IV [a], Mayuko Nakamaru [c], Hisashi Ohtsuki [c,d]

[a] Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA
[b] Berkman Center for Internet and Society, Harvard University, Cambridge, MA 02138, USA
[c] Department of Value and Decision Science, Tokyo Institute of Technology, Tokyo 152-8552, Japan
[d] PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

ARTICLE INFO

Article history:
Received 1 December 2009
Received in revised form
4 June 2010
Accepted 4 June 2010
Available online 9 June 2010

Keywords:
Evolutionary game theory
Prisoner's dilemma
Public goods game
Structured populations
Spite

ABSTRACT

The evolution of cooperation is one of the great puzzles in evolutionary biology. Punishment has been suggested as one solution to this problem. Here punishment is generally defined as incurring a cost to inflict harm on a wrong-doer. In the presence of punishers, cooperators can gain higher payoffs than non-cooperators. Therefore cooperation may evolve as long as punishment is prevalent in the population. Theoretical models have revealed that spatial structure can favor the co-evolution of punishment and cooperation, by allowing individuals to only play and compete with those in their immediate neighborhood. However, those models have usually assumed that punishment is always targeted at non-cooperators. In light of recent empirical evidence of punishment targeted at cooperators, we relax this assumption and study the effect of so-called 'anti-social punishment'. We find that evolution can favor anti-social punishment, and that when anti-social punishment is possible costly punishment no longer promotes cooperation. As there is no reason to assume that cooperators cannot be the target of punishment during evolution, our results demonstrate serious restrictions on the ability of costly punishment to allow the evolution of cooperation in spatially structured populations. Our results also help to make sense of the empirical observation that defectors will sometimes pay to punish cooperators.

1. Introduction

Explaining the evolution of cooperation is an issue of central importance in evolutionary biology as well as the social sciences. Why does the competitive process of natural selection often lead to altruistic cooperation, in which individuals pay a cost to give a benefit to others? To answer this question, numerous mechanisms for the evolution of cooperation have been proposed (Nowak, 2006; Lehmann and Keller, 2006; West et al., 2007), including kin selection (Hamilton, 1964; Frank, 1998), direct reciprocity (Trivers, 1971) and indirect reciprocity (Alexander, 1987; Nowak and Sigmund, 2005).

In direct reciprocity, my actions towards you depend on what you have done to me in the past. Axelrod (1984) found that the tit-for-tat (TFT) strategy was a winning strategy in his computer tournament. A TFT player cooperates in the first round of the repeated interaction. After the first round, a TFT player takes whatever action his opponent took in the previous round. For example, an ALLD player, who always defects in a repeated interaction, receives cooperation from a TFT player in the first round. However, the ALLD player is 'punished' by TFT with mutual defection in subsequent rounds. Therefore, cheating does not pay if the repeated interaction lasts long enough, and the TFT strategy can be evolutionarily stable against the ALLD strategy (Axelrod and Hamilton, 1981). Numerous behavioral experiments have demonstrated that direct reciprocity can lead to stable cooperation in repeated games (Wedekind and Milinski, 1996; Dal Bó, 2005; Dreber et al., 2008; Dal Bó and Frechette, 2009; Rand et al., 2009a; Fudenberg et al., 2010).

Another possible way to enforce cooperation is to rely on more explicit punishment (Clutton-Brock and Parker, 1995a,b; Sigmund, 2007; Rand et al., 2009b). While TFT 'punishes' non-cooperative partners by responding to selfishness with reciprocal selfishness, explicit punishment strategies choose to inflict harm on non-cooperative partners, often at a cost to the punisher. For example, in paper wasps, subordinates that cheat by signaling an inflated status of dominance receive more aggression from other wasps (Tibbetts and Dale, 2004). Similarly, monkeys that do not share food are often punished (Hauser, 1992). Punishment is also

* Corresponding author at: Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA.
E-mail address: drand@fas.harvard.edu (D.G. Rand).

0022-5193/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2010.06.010

known as 'policing' in studies of social insects, where queens or workers sometimes attack other nestmate workers that attempt to produce offspring by themselves (Ratnieks et al., 2006; Wenseleers and Ratnieks, 2006). In behavioral experiments with humans, punishment has been shown to stabilize cooperation in multi-player social dilemmas (Yamagishi, 1986; Ostrom et al., 1992; Fehr and Gächter, 2000, 2002), leading to higher payoffs after an initial learning period (Gächter et al., 2008; Rand et al., 2009b). As a proximate explanation of such pro-social punishment, the desire for egalitarian outcomes (Fehr and Schmidt, 1999; Johnson et al., 2009) and/or anger directed at norm violators (Fehr and Gächter, 2002) have been recently suggested.

Under the threat of punishment, a cooperative strategy can perform better than a non-cooperative strategy, because the latter suffers the costs of being punished. Therefore cooperation may evolve in populations where punishment is prevalent. When punishment is also costly to punisher, however, we face another puzzle: how does punishment evolve? Cooperators who do not participate in punishment avoid the cost of punishment. Thus, they out-compete those who cooperate and punish. This puzzle is well known as the 'second-order free rider problem' (Oliver, 1980; Yamagishi, 1986; Boyd and Richerson, 1992; Panchanathan and Boyd, 2004).

Such 'costly punishment' (Fehr and Gächter, 2000, 2002) has been shown to co-evolve with cooperation when, for example, cooperation and punishment are such strongly linked traits that cooperators almost always participate in punishment (Axelrod, 1986; Boyd and Richerson, 1992; Nakamaru and Iwasa, 2005; Lehmann et al., 2007). If cooperators punish defectors, the advantage of selfish individuals is greatly reduced, and cooperation can co-evolve with punishment through group selection (Boyd et al., 2003), although the possibility of retaliation is problematic for cooperation (Janssen and Bushman, 2008). Cooperation and punishment can co-evolve when those who punish defectors are compensated by being more likely to receive cooperation from others (Gardner and West, 2004). Recent models of cultural evolution suggest that weak conformist transmission can stabilize punishment and hence promotes the evolution of cooperation (Henrich and Boyd, 2001; Henrich, 2004). The option to abstain from the game also favors the co-evolution of cooperation and punishment (Fowler, 2005; Hauert et al., 2007; Traulsen et al., 2009). In repeated two-player games, however, it has been shown that costly punishment does not promote cooperation, as traditional tit-for-tat style direct reciprocity is sufficient (Rand et al., 2009b). In contrast to such repeated games, a credible threat of costly punishment could be effective in one-shot games, for example in the situation where resources are so scarce that a long-term reciprocal interaction, and therefore use of the TFT strategy, is unlikely.

Nakamaru and Iwasa (2005, 2006) have shown that costly punishment promotes cooperation in a two-stage game played in a spatially structured population. Nakamaru and Iwasa (2005) studied two strategies: an altruist-punisher, who cooperates with others and punishes non-cooperators, and a selfish-non-punisher, who neither cooperates with others nor punishes non-cooperators. In a subsequent paper (Nakamaru and Iwasa, 2006), they included two more strategies in their analysis: an altruist-non-punisher, who cooperates but does not punish, and a selfish-punisher, who does not cooperate but punishes non-cooperators (see also Sigmund et al., 2001).

These theoretical models have studied the effect of allowing players to punish non-cooperators. In addition to such pro-social punishment, however, numerous behavioral experiments have found that a significant fraction of non-cooperators will pay to punish cooperators (Shinada et al., 2004; Denant-Boemont et al., 2007; Dreber et al., 2008; Herrmann et al., 2008; Nikiforakis, 2008; Gächter and Herrmann, 2009; Wu et al., 2009). For example, a series of cross-cultural public goods game experiments found a great deal of cross-cultural variation in the extent to which cooperators versus non-cooperators were targeted with punishment (Herrmann et al., 2008). In the most extreme cases, participants in countries such as Greece and Oman were as likely to punish those who contributed more than them as those who contributed less. The high level of punishment directed at cooperators in this and other experiments indicates that this behavior is not merely the result of errors or lack of comprehension, but is instead a surprising aspect of human behavior requiring explanation. Throughout this paper, we refer to this punishment targeted at cooperators as 'anti-social punishment', in order to distinguish it from the (usual) punishment that is targeted at non-cooperators. Anti-social punishment runs counter to the common assumptions about why people choose to punish (Johnson et al., 2009; Fehr and Gächter, 2000, 2002). These strategies that pay a cost for cooperators to incur a cost are often excluded from evolutionary models of cooperation and punishment. In addition to the empirical evidence for anti-social punishment, it seems most appropriate for evolutionary models to allow the full set of possible strategies (of a given complexity) and to ask what strategies emerge via natural selection, rather than restricting the strategy set to only include particularly attractive strategies.

In this paper, we ask how including anti-social punishment affects the evolution of cooperation. Is this 'dark side' of punishment in behavioral experiments just irrational, erroneous behavior, or can it in fact be favored by natural selection? Does punishment still promote cooperation when anti-social punishment is possible? To explore these questions, we adopt the framework used by Nakamaru and Iwasa (2005, 2006) and study the consequences of including anti-social punishment.

2. Strategies and payoffs

We consider a two-stage game with cooperation followed by punishment (as in Boyd and Richerson, 1992; Sigmund et al., 2001; Nakamaru and Iwasa, 2005, 2006; Boyd et al., 2003; Brandt et al., 2003, 2006; Fowler, 2005; Hauert et al., 2007; Traulsen et al., 2009). In the first stage, each player has two choices, defection (D) or cooperation (C). Defection means doing nothing, such that all players receive zero payoffs. Cooperation means paying a cost c for another to get a benefit b ($b>c>0$). In the second stage, each player can choose to punish each other player or not, conditioned on the other person's action in the first stage. By punishing, a player pays a cost α for the other person to incur a loss β ($\alpha,\beta>0$). Withholding punishment results in both players receiving zero payoffs in the second stage. We do not consider mixed strategies in the following analysis, and restrict our attention to pure reactive strategies.

Since the action in the second stage is conditioned on the other players' action in the first stage, there are four possible strategies in the second stage. A non-punisher (N) punishes no one. A pro-social punisher (P) punishes defectors only. An anti-social punisher (A) punishes cooperators only. A spiteful punisher (S) punishes both cooperators and defectors. A combination of the action in the first stage (D,C) and the strategy in the second stage (N, P, A, S) defines one's strategy in the game. Thus we have eight possible strategies: DN, DP, DA, DS, CN, CP, CA and CS. For example, a CP-strategist cooperates in the first stage and punishes defectors in the second stage (therefore deemed as a 'strong reciprocator' (Fehr and Fischbacher, 2003; Bowles and Gintis, 2004)). Table 1 summarizes these eight strategies.

We can classify these eight strategies into two categories: strategies which do not punish other players using the same

Table 1
The eight strategies considered in the present paper.

Strategy	Cooperate?	Harm cooperators?	Harm defectors?
DN	No	No	No
DP	No	No	Yes
DA	No	Yes	No
DS	No	Yes	Yes
CN	Yes	No	No
CP	Yes	No	Yes
CA	Yes	Yes	No
CS	Yes	Yes	Yes

strategy (hereafter 'self-consistent'), and strategies which do punish other players using the same strategy (hereafter 'self inconsistent'). The strategies DN, DA, CN and CP are self-consistent. Conversely, the strategies DP, DS, CA and CS are self-inconsistent.

The 8×8 payoff matrix of the pairwise game is given in Eq. (1). As can be seen, the strategies DN and DA are Nash equilibria for all payoff values, and CP is a Nash equilibrium when $\beta > c$.

$$\begin{array}{c} \\ \text{DN} \\ \text{DP} \\ \text{DA} \\ \text{DS} \\ \text{CN} \\ \text{CP} \\ \text{CA} \\ \text{CS} \end{array} \begin{array}{c} \begin{array}{cccccccc} \text{DN} & \text{DP} & \text{DA} & \text{DS} & \text{CN} & \text{CP} & \text{CA} & \text{CS} \end{array} \\ \begin{pmatrix} 0 & -\beta & 0 & -\beta & b & b-\beta & b & b-\beta \\ -\alpha & -\alpha-\beta & -\alpha & -\alpha-\beta & b & b-\beta & b & b-\beta \\ 0 & -\beta & 0 & -\beta & b-\alpha & b-\alpha-\beta & b-\alpha & b-\alpha-\beta \\ -\alpha & -\alpha-\beta & -\alpha & -\alpha-\beta & b-\alpha & b-\alpha-\beta & b-\alpha & b-\alpha-\beta \\ -c & -c & -c-\beta & -c-\beta & -c+b & -c+b & -c+b-\beta & -c+b-\beta \\ -c-\alpha & -c-\alpha & -c-\alpha-\beta & -c-\alpha-\beta & -c+b & -c+b & -c+b-\beta & -c+b-\beta \\ -c & -c & -c-\beta & -c-\beta & -c+b-\alpha & -c+b-\alpha & -c+b-\alpha-\beta & -c+b-\alpha-\beta \\ -c-\alpha & -c-\alpha & -c-\alpha-\beta & -c-\alpha-\beta & -c+b-\alpha & -c+b-\alpha & -c+b-\alpha-\beta & -c+b-\alpha-\beta \end{pmatrix} \end{array}. \tag{1}$$

For reference, consider the sub-model where punishment cannot be targeted at cooperators. In this case, the strategies A and S are not allowed in the second stage, so the payoff matrix (1) is reduced to

$$\begin{array}{c} \\ \text{DN} \\ \text{DP} \\ \text{CN} \\ \text{CP} \end{array} \begin{array}{c} \begin{array}{cccc} \text{DN} & \text{DP} & \text{CN} & \text{CP} \end{array} \\ \begin{pmatrix} 0 & -\beta & b & b-\beta \\ -\alpha & -\alpha-\beta & b & b-\beta \\ -c & -c & -c+b & -c+b \\ -c-\alpha & -c-\alpha & -c+b & -c+b \end{pmatrix} \end{array}. \tag{2}$$

In the following analysis, we will concentrate on the 'score-dependent viability model' (Nakamaru et al., 1997; Irwin and Taylor, 2001; Nakamaru and Iwasa, 2005, 2006; Sekiguchi and Nakamaru, 2009) (hereafter 'viability model'). We chose to study the viability model as it has previously been shown to be very effective in promoting the co-evolution of punishment and cooperation in structured populations (Nakamaru and Iwasa, 2005, 2006), and therefore presents the greatest challenge to anti-social punishment. We will compare evolutionary dynamics in unstructured populations and lattice structured populations, and investigate the effect on the evolution of cooperation of introducing anti-social punishment into the game. We do so by expanding the strategy set to include all combinatorially possible pure strategies (shown in Eq. (1)), as opposed to making any structural changes to the game (for example creating the opportunity to retaliate as in Janssen and Bushman, 2008) or increasing the complexity of the possible strategies.

3. *z*-mixed population model

First we study viability game dynamics in populations with no spatial structure. This viability updating rule can be interpreted as either genetic evolution or social learning. In the context of genetic evolution, an organism's probability of death is affected by its payoff, while reproduction is random. In the context of social learning, a person's probability of abandoning her current strategy is determined by her payoff, after which she randomly picks a new strategy from the population to imitate (a type of conformism).

Groups of size $z+1$ are randomly selected from the population, and each player interacts with her z other group members to obtain a game-payoff, f. In each generation a random player is given a chance to update her strategy (i.e. has some chance of death). With probability

$$d(f) = \gamma \exp[-\theta f], \tag{3}$$

she abandons her current strategy (i.e. dies) and randomly adopts the strategy of one of the z players she just interacted with (i.e. one of the z players is randomly chosen to reproduce). This model was called the complete mixing model' by Nakamaru et al. (1997) and Nakamaru and Iwasa (2005, 2006). However, as we will see below, this model is different from a usual 'well-mixed population' model often assumed in evolutionary game theory. To avoid potential confusion, we call the above model 'z-mixed population model', where z represents the number of players one interacts with. Unlike the traditional well-mixed population model, agents in the z-mixed population model interact with, and compete with, only a subset of the population. The $z\to\infty$ limit recovers the traditional well-mixed population dynamics.

Here $\gamma,\theta > 0$ are constants which influence the intensity of selection. The intuitive consequence of Eq. (3) is that players with lower payoffs are more likely to change their strategies.

3.1. When anti-social punishment is not allowed

First we consider the four-strategy game, Eq. (2). It has previously been shown that in a z-mixed population, the strategies CP and CN are selected when β is large, regardless of initial frequencies of strategies (Nakamaru and Iwasa, 2006). This result is surprising given that DN is always a Nash equilibrium strategy in the game, Eq. (2). The key lies in the assumptions of the viability model. In the viability model, (i) one interacts with only a subset of the population (finite z), and (ii) those whom one mimics are the same as those whom one interacts with. Such a local population regulation significantly affects stability conditions of strategies, and is a common feature of many models of evolution in structured populations (Nowak and May, 1992; Wilson et al., 1992; Taylor, 1992; Killingback and Doebeli, 1996; Szabo and Toke, 1998; Hauert and Doebeli, 2004; Ohtsuki et al.,

2006; Santos et al., 2006; Fu et al., 2007). In particular, Nash equilibrium strategies may no longer be stable strategies in the viability model, and non-Nash equilibrium strategies may be stable.

To see how viability updating changes the stability condition, imagine the game between two arbitrary strategies, X and Y

$$\begin{array}{c} \\ \mathrm{X} \\ \mathrm{Y} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{X} & \mathrm{Y} \end{array} \\ \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \end{array}. \tag{4}$$

The condition for strategy X to be a strict Nash equilibrium is $a_1 > a_3$, namely, X performs better than Y against X. However, in the viability model it can be shown that strategy X is stable against invasion by strategy Y when

$$(z-1)a_1 + a_2 > z a_3. \tag{5}$$

Here z is the number of others that one interacts with. See Appendix A for its derivation. Interestingly, a_2, which is the payoff of X playing against Y, appears in the stability condition. The intuitive explanation for this is that a Y-player affects the payoffs of those X players she then mimics. As z increases, the effect of a_2 diminishes, but never disappears as long as z is finite.

Using the stability condition, Eq. (5), we obtain for general z that CP is robust against invasion by DP and DN when

$$\beta > c + \frac{b+\alpha}{z}. \tag{6}$$

Note that CP is always neutral to CN, because there are no defectors to punish in the population of only CP and CN.

We can further extend previous analysis by introducing errors. We assume that there is a (sufficiently) small chance ε of error in execution of intended actions (Molander, 1985; May, 1987; Nowak and Sigmund, 1989, 1992, 1993; Fudenberg and Maskin, 1990; Fudenberg and Tirole, 1991; Lindgren, 1991; Lindgren and Nordahl, 1994; Boerlijst et al., 1997; Wahl and Nowak, 1999; Brandt and Sigmund, 2006; Fudenberg et al. 2010). More precisely, one fails to perform an intended action in the first stage and/or the second stage independently with probability ε. In the first stage, with probability ε one mistakenly chooses D though her intended action is C, and vice versa. In the second stage, with probability ε one fails to punish others though she intended to, and vice versa.

Such errors resolve the neutrality between CN and CP. In the presence of small errors, the payoff matrix between CN and CP becomes

$$\begin{array}{c} \\ \mathrm{CN} \\ \mathrm{CP} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{CN} & \mathrm{CP} \end{array} \\ \begin{pmatrix} b-c+(-b+c-\alpha-\beta)\varepsilon & b-c+(-b+c-\alpha-2\beta)\varepsilon+2\beta\varepsilon^2 \\ b-c+(-b+c-2\alpha-\beta)\varepsilon+2\alpha\varepsilon^2 & b-c+(-b+c-2\alpha-2\beta)\varepsilon+(2\alpha+2\beta)\varepsilon^2 \end{pmatrix} \end{array}. \tag{7}$$

Since ε is sufficiently small, it has an effect only when payoffs in Eq. (2) are equal (i.e. when two strategies are neutral). Neglecting ε^2 terms and using Eq. (5), we obtain that CP is stable against CN when

$$\beta > z\alpha. \tag{8}$$

Combining (6) and (8) leads to the ESS condition of strategy CP, as

$$\beta > \max\left\{c + \frac{b+\alpha}{z}, z\alpha\right\}. \tag{9}$$

Similarly, the ESS condition for the other three strategies can be derived. Fig. 1A shows the parameter regions where each strategy is an ESS. The results are equivalent to those found in the absence of error (Nakamaru and Iwasa, 2006), except that CN is never an ESS. In particular, we see that the cooperative strategy CP is the only ESS when β is sufficiently large.

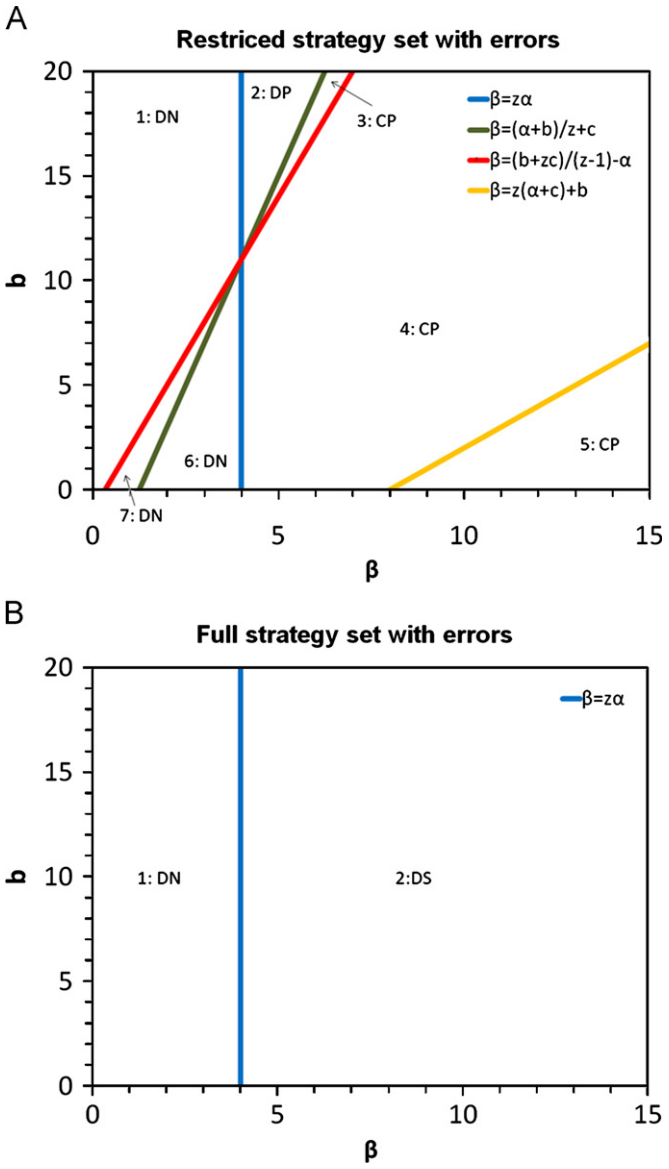

Fig. 1. Evolutionary stability under the z-mixed population model. The strategies that are ESS are indicated in each region of the (b,β) parameter space. We set $c=1$ and $\alpha=1$. We consider small errors. (A) When anti-social punishment is not allowed, CP is ESS in regions 3, 4 and 5. (B) When anti-social punishment is allowed, CP is never ESS. Either DN or DS is the unique ESS.

3.2. When anti-social punishment is allowed

We now allow the option of anti-social punishment, making the punishment strategies A (anti-social punisher) and S (spiteful strategy) possible in the second stage of the game. The game is thus described by the full 8×8 matrix given by Eq. (1).

In the absence of errors ($\varepsilon=0$), we find three pairs of strategies that are neutral in addition to CN–CP. DN and DA are neutral because there are no cooperators to punish. DP and DS are neutral because they always punish themselves and each other. CA and CS are neutral because they also always punish themselves and each other. With the introduction of errors, the payoff matrices for these pairs change to

$$\begin{array}{c} \\ \mathrm{DN} \\ \mathrm{DA} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{DN} & \mathrm{DA} \end{array} \\ \begin{pmatrix} (b-c-\alpha-\beta)\varepsilon & (b-c-\alpha-2\beta)\varepsilon+2\beta\varepsilon^2 \\ (b-c-2\alpha-\beta)\varepsilon+2\alpha\varepsilon^2 & (b-c-2\alpha-2\beta)\varepsilon+(2\alpha+2\beta)\varepsilon^2 \end{pmatrix} \end{array}, \tag{10}$$

$$\begin{array}{c} \\ \text{DP} \\ \text{DS} \end{array} \begin{array}{c} \begin{array}{cc} \text{DP} & \text{DS} \end{array} \\ \begin{pmatrix} -\alpha-\beta+(b-c+2\alpha+2\beta)\varepsilon+(-2\alpha-2\beta)\varepsilon^2 & -\alpha-\beta+(b-c+2\alpha+\beta)\varepsilon-2\alpha\varepsilon^2 \\ -\alpha-\beta+(b-c+\alpha+2\beta)\varepsilon-2\beta\varepsilon^2 & -\alpha-\beta+(b-c+\alpha+\beta)\varepsilon \end{pmatrix} \end{array}, \tag{11}$$

$$\begin{array}{c} \\ \text{CA} \\ \text{CS} \end{array} \begin{array}{c} \begin{array}{cc} \text{CA} & \text{CS} \end{array} \\ \begin{pmatrix} b-c-\alpha-\beta+(-b+c+2\alpha+2\beta)\varepsilon+(-2\alpha-2\beta)\varepsilon^2 & b-c-\alpha-\beta+(-b+c+2\alpha+\beta)\varepsilon-2\alpha\varepsilon^2 \\ b-c-\alpha-\beta+(-b+c+\alpha+2\beta)\varepsilon-2\beta\varepsilon^2 & b-c-\alpha-\beta+(-b+c+\alpha+\beta)\varepsilon \end{pmatrix} \end{array}. \tag{12}$$

Taking into account Eqs. (7, 10–12), the evolutionary stability of each strategy in the game Eq. (1) under the z-mixed population model is derived.

(i) DN is ESS if $\beta < z\alpha$, and is invaded by DP, DA, and DS if $\beta > z\alpha$.
(ii) DP is invaded by DN and DA if $\beta < z\alpha$, and invaded by DS if $\beta > z\alpha$. Thus DP is never ESS.
(iii) DA is invaded by DN if $\beta < z\alpha$, and invaded by DP and DS if $\beta > z\alpha$. Thus DA is never ESS.
(iv) DS is invaded by DN, DP and DA if $\beta < z\alpha$, and is ESS if $\beta > z\alpha$.
(v) CN is always invaded by DN and DP. Thus CN is never ESS.
(vi) CP is invaded by CN if $\beta < z\alpha$, and invaded by CA and CS if $\beta > z\alpha$. Thus CP is never ESS.
(vii) CA is always invaded by DN and DP. Thus CA is never ESS.
(viii) CS is always invaded by DA and DS. Thus CS is never ESS.

Fig. 1B summarizes the result. If $\beta < z\alpha$, DN is the unique ESS. If $\beta > z\alpha$, DS is the unique ESS. It is noteworthy that CP is never an ESS under the full strategy set. Anti-social behavior and spite destroy cooperation. As we saw above, if $\beta > z\alpha$, CP is not invaded by CN, DN, or DP and therefore it can be ESS in the absence of anti-social punishment. When anti-social punishment is available, however, CA and CS can invade CP if $\beta > z\alpha$. The reason for this invasion is simple. Imagine a CA (or CS)-player in a population of CP. The CA player punishes z CP players, so he pays the total cost of $z\alpha$, whereas each CP player who meets the CA player incurs the loss β. If this loss is greater than the total cost the CA pays, CP is more likely to change to CA than vice versa, and therefore CA propagates in the population of CP. CA players maximize their relative fitness in a population of CP players by avoiding the punishment they would receive for defecting and by harming others to enhance the probability that they are mimicked. However, invasion by CA or CS players does not suggest maintenance of cooperation by anti-social punishment, as they are easily invaded by various defecting strategies. Thus evolution always leads to DS if $\beta < z\alpha$.

To summarize, when punishment is sufficiently effective (i.e. $\beta > z\alpha$), it is adaptive to punish others and increase one's relative payoff no matter what strategy the others adopt in the first stage (see Appendix B). Hence the optimal target of punishment should be not only defectors but also cooperators, and the condition $\beta > z\alpha$ is simply the condition under which unconditional punishment is adaptive in the z-mixed population model. When it was only possible to punish defectors, this power of punishment allowed CP to invade DN. Once anti-social punishment is available, however, cooperators no longer have an advantage, and spiteful defection, DS, fares best. The $\beta > z\alpha$ condition reveals that the success of spiteful punishers relies on agents interacting and competing with a limited number of others. In the $z \to \infty$ limit of a well-mixed infinite population, DN is the unique ESS for all payoff values.

4. Lattice model

We now consider a structured population in which players are arranged on a square lattice. Each player interacts with the four players in her von Neumann neighborhood and obtains a game-payoff. Thus this lattice model is the structured counterpart to the z-mixed population model with $z=4$, and is equivalent to a series of overlapping linear public goods games (e.g. Santos et al., 2008), where each player's cooperative actions affect her four neighbors. Other features of the model are the same as the z-mixed population model as described in the previous section. Agent based simulations are used to explore the evolutionary dynamics in this structured population.

4.1. When anti-social punishment is not allowed

Examining the restricted four-strategy game on a lattice, we find similar results to the z-mixed population model. As Fig. 2 shows, the defecting strategy, DN, is favored when β is small, and the cooperative strategies, CN and CP, are favored when β is large. Thus our inclusion of ε errors does not change results on the lattice compared to what has been found previously (Nakamaru and Iwasa, 2006). In the absence of anti-social punishment, there is a large parameter region in which cooperative strategies dominate on the lattice, regardless of initial conditions.

4.2. When anti-social punishment is allowed

We now explore the dynamics on the lattice when strategies which punish cooperators are included, Eq. (1). Simulation results are shown in Fig. 3. Results are dependent on initial strategy frequencies. When DN or CP is most abundant initially, we see a

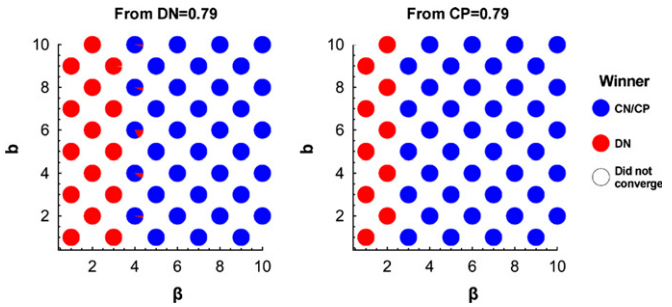

Fig. 2. Evolutionary outcomes of the restricted strategy set without anti-social punishment on a regular square lattice. When β is small, DN wins regardless of initial frequencies of strategies; when β is large, CP wins regardless of initial frequencies of strategies. Starting from the specified initial frequency, 50 agent based simulations are run and the winning strategy is recorded. The blue portion of each circle indicates the fraction of runs where CP wins or CP and CN coexist; the red portion, the fraction of runs where DN wins; the white portion, the fraction of runs in which there was no convergence after 125,000,000 generations. We explore the (b,β) parameter space, setting $c=1$ and $\alpha=1$. We consider small errors, $\varepsilon=0.01$, and a 50×50 lattice for a total population size $N=2500$. We use viability updating with parameter values $\gamma=0.1$ and $\theta=0.1$. (A) Initial frequency of strategies: DN=0.79, DP=0.07, CN=0.07, CP=0.07. (B) Initial frequency of strategies: DN=0.07, DP=0.07, CN=0.07, CP=0.79. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

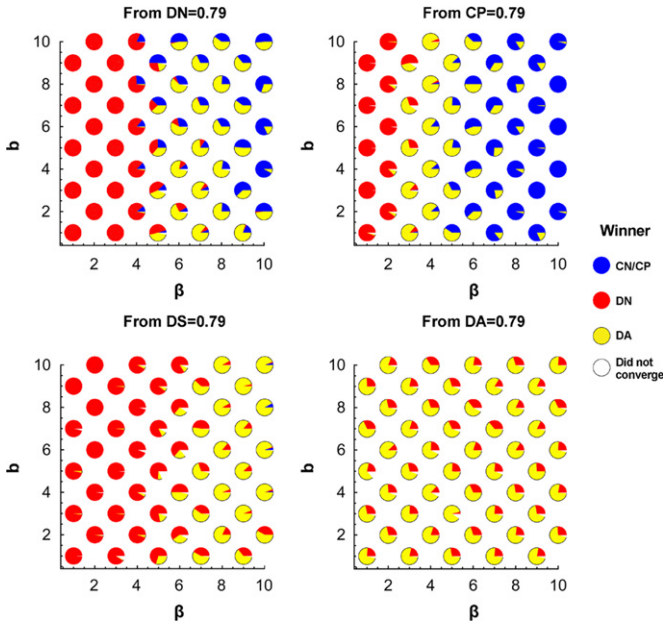

Fig. 3. Evolutionary outcomes of the full strategy set on a regular lattice. Cooperation cannot invade a population of antisocial defectors, DA. Starting from the specified initial condition, 50 agent based simulations are run and the winning strategy recorded. The blue portion of each circle indicates the fraction of runs where CP wins or CP, CN, CA and/or CS coexist; the red portion, the fraction of runs where DN wins; the yellow portion the fraction of runs where DA wins; the white portion, the fraction of runs in which there was no convergence after 125,000,000 generations. DP and DS never win. We consider small errors, $\varepsilon = 0.01$, and a 50×50 lattice for a total population size $N = 2500$. We use viability updating with parameter values $\gamma = 0.1$ and $\theta = 0.1$. We explore the (b,β) parameter space, setting $c = 1$ and $\alpha = 1$. Additional simulations find qualitatively similar results for $\alpha = 0.5$ and $\alpha = 1.5$. Initial population density (A) DN=0.79, all others 0.03; (B) CP=0.79, all others 0.03; (C) DS=0.79, all others 0.03; and (D) DA=0.79, all others 0.03. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

similar pattern: the majority of the time DN wins when β is small, DA wins when β is intermediate, and CP wins when β is large. When DS is most abundant initially, DN again wins when β is small, DA wins when β is large, and further simulations find that CP wins when $\beta > 10$. When DA is most abundant initially, however, the outcome is very different. Regardless of b or β, DA wins in the majority of cases. DN wins occasionally. Cooperation never invades a resident population of anti-social defectors.

These results are qualitatively different from what we saw in the z-mixed population. First, DS never wins on the lattice, whereas DS was the sole ESS in the z-mixed population model when β was large. A possible explanation for this difference involves spatial correlations. Since offspring disperse locally on the lattice, DS players are very likely to interact with other DS players (i.e. their 'relatives'). Because DS is self-inconsistent in the sense that one DS player harms other DS-players, the target of the DS players' spiteful punishment tends to be other DS players. Therefore DS, as well as the other self-inconsistent strategies DP, CA and CS, cannot propagate on the lattice (see also Appendix B). A pairwise invasion analysis based on computer simulations (Table 2) shows that across β values, each self-inconsistent strategy is invaded by the first-round equivalent self-consistent strategies (i.e. DP and DS invaded by DN and DA, CA and CS invaded by CN and CP).

The second major difference we observe is that on the lattice, anti-social defectors (DA) win in a large region of the parameter space, whereas DA is never an ESS in the z-mixed population. This is because DP and DS, which are potential invaders of DA in the z-mixed population model, are self-inconsistent and thus do not perform well on the lattice as discussed above. In contrast, DA is self-consistent and punishes only cooperators, which protects DA from invasion by CP. The pairwise invasion analysis in Table 2 suggests that strategies other than DN cannot invade DA on the lattice. DN can occasionally invade DA, but only due to a small advantage from the ε tiny error probability (see Eq. (10)). Hence DN and DA are almost neutral. In the z-mixed population model DA can invade DN when β is large. In contrast DA never invades DN on the lattice. Again this is because in the presence of errors, DA engages in self-inconsistent punishment of erroneous cooperation: in z-well mixed populations, this punishment gives DA an advantage over DN; but on the lattice, the recipients of DA's punishment are most likely also DA, and so DA is at a disadvantage.

Overall, the lattice simulation results using the full strategy set are quite different from what occurs in the restricted strategy set where anti-social punishment is not possible. When anti-social punishment is excluded, there is a wide parameter region where cooperation wins regardless of initial conditions (Fig. 2). However, the introduction of anti-social punishment eliminates this region, hinders cooperation, and anti-social defectors prevail (Fig. 3).

Table 2
Pairwise invasion analysis in a lattice structured population.

	Resident							
	DN	DP	DA	DS	CN	CP	CA	CS
$\beta = 1$								
Invader								
DN		1	0.24	1	1	1	1	1
DP	0		0	0.35	0	0	1	1
DA	0	1		1	1	1	1	1
DS	0	0	0		0	0	1	1
CN	0	0	0	0		0	1	1
CP	0	0	0	0	0		1	1
CA	0	0	0	0	0	0		0
CS	0	0	0	0	0	0	0	
$\beta = 5$								
Invader								
DN		1	0.20	1	1	0	1	1
DP	0		0	0.05	0	0	0	0
DA	0	1		1	1	0.17	1	1
DS	0	0	0		0	0	1	1
CN	0	1	0	0		0	1	1
CP	0.95	1	0	1	0		1	1
CA	0	0	0	0	0	0		0.30
CS	0	0	0	0	0	0	0	
$\beta = 10$								
Invader								
DN		1	0.21	1	1	0	1	1
DP	0		0	0	0	0	0	0
DA	0	1		1	1	0	1	1
DS	0	0.03	0		0	0	1	0
CN	0	1	0	0		0	1	1
CP	1	1	0	1	0		1	1
CA	0	0	0	0	0	0		0.08
CS	0	1	0	0	0	0	0	

In each cell we compare two strategies, resident and mutant. The resident's initial frequency is 0.97, and the invader's initial frequency is 0.03. The fraction of simulations in which the invader takes over the whole population is shown. We set $b = 5$, $c = 1$ and $\alpha = 1$, and conduct 50 simulations, each lasting 125,000,000 generations. We consider small errors, $\varepsilon = 0.01$, and a 50×50 lattice for a total population size $N = 2500$. We use viability updating with parameter values $\gamma = 0.1$ and $\theta = 0.1$.

5. Discussion

Here we have analyzed the effect of anti-social punishment on the evolution of cooperation. We have studied score-dependent

viability dynamics (Nakamaru and Iwasa, 2005, 2006). We include all eight possible pure reactive strategies as well as tiny behavioral errors, and compare the z-mixed population model and the spatially structured lattice model.

Results in the z-mixed population model clearly demonstrate that the inclusion of anti-social punishment destroys the evolutionary success of cooperation. We found that cooperation is never ESS under the full strategy set. Spiteful defectors who always harm others (DS) are the only ESS if the punishment technology is sufficiently efficient (sufficiently large effect to cost ratio). Our lattice model analysis confirms the difficulties anti-social punishment poses for cooperation. On the lattice, cooperators can never invade a population of anti-social defectors who punish cooperators (DA). This anti-social defection strategy is highly successful across the wide parameter ranges we tested. These results show that anti-social punishment can be favored by natural selection, and highlight the dangers associated with allowing the option for costly punishment. Not only is it possible for cooperators to punish defectors, but also the opposite can occur. Together with the previous results of Nakamaru and Iwasa (2006), our results suggest that punishment only successfully promotes cooperation in structured populations if it is only possible to harm defectors but not cooperators. Our results suggest that there is an evolutionary imperative for defectors to seek out ways to punish cooperators, and thus models which exclude this phenomenon may give skewed results favoring the evolution of cooperation.

At first it seems counter-intuitive that successful self-interested players would pay a cost to harm others in the second (and final) stage of a two-stage game. The explanation lies in the nature of evolutionary competition in settings where you learn from the same players you interact with. Here it is not one's absolute payoff that matters, but rather how much you have relative to others (Page et al., 2000). Although punishers (P), anti-social punishers (A), and spiteful punishers (S) in our model seem to perform spite at a level of absolute payoffs because both punisher and punishee incur immediate costs, the actual effect of these behaviors is to enhance the punisher's probability of replacing the victims of punishment: inflicting harm which reduces the others' payoff by more than it reduces yours improves your relative payoff. Therefore, spiteful strategies which harm others can be at an advantage in a spatially structured setting, as long as they do not preferentially harm agents with their own strategy (Hamilton, 1970; Wilson, 1975; Nakamaru et al., 1997; Foster et al., 2000; Nakamaru and Iwasa, 2005; Lehmann et al., 2006; El Mouden and Gardner, 2008). In this sense, punishment in our model does not represent genuine spite but can be classified as a selfish behavior (West and Gardner, 2010). Conversely, evolution has been shown to disfavor costly punishment in repeated games played in well-mixed populations (Rand et al., 2009b). The effect of costly punishment in repeated games with local interaction and competition, however, is an interesting question for future study, as is the role of anti-social punishment in games with continuous (as opposed to binary) traits (Nakamaru and Dieckmann, 2009). Furthermore, while our analysis has demonstrated that the population structure considered in this paper can favor anti-social punishment, a more explicitly structured population, such as a population subdivided into many small groups, may not, as the balance between within-group selection and between-group selection could change (Wilson, 1975; Okasha, 2006; Traulsen and Nowak, 2006; Wilson and Wilson, 2007). Similarly, incorporating explicit inter-group conflicts into our model could increase the importance of group-level selection, and therefore could change the relative importance of anti-social and pro-social punishment (Bowles, 2009; Lehmann and Feldman, 2008). These issues deserve further study.

We have framed the cooperative dilemma faced by players in our model as a series of pairwise interactions. However, because players are unconditional cooperators or defectors, the game we study is exactly equivalent to a z+1-player public goods game followed by punishment (Boyd and Richerson, 1992; Fehr and Gächter, 2000, 2002; Sigmund et al., 2001; Nakamaru and Iwasa, 2005, 2006; Boyd et al., 2003; Brandt et al., 2003, 2006; Fowler, 2005; Hauert et al., 2007; Traulsen et al., 2009). In the z-well-mixed population, a cooperator pays a cost to benefit each other member of her group, which is drawn randomly each round; this corresponds to a public goods game played under the 'stranger matching' protocol (Fehr and Gächter, 2000, 2002). In the lattice population structure, a cooperator pays a cost to benefit her $z=4$ neighbors; this corresponds to a series of overlapping public goods games with fixed group compositions. Thus the results we observe regarding anti-social punishment are not unique to pairwise interactions, but rather apply to collective action problems confronted by groups of individuals.

In our modeling framework, agents update their strategies through an evolutionary process. As opposed to prospectively calculating the strategy that would maximize one's payoff, our agents are motivated to change strategy when their payoffs are low, and then imitate the behavior of others they observe. Thus strategies with higher payoffs tend to become more common in the population. If instead our agents picked strategies based on rational self-interest using Eq. (1) with ε errors, they would never cooperate either with or without anti-social punishment as DN is the unique Nash equilibrium. Thus in both extremes of strict imitation and strict rationality, cooperation is not stable using the full strategy set. Intermediate cases in which agents use some combination of imitation and prospective reasoning are an interesting subject for further study.

The analysis presented here raises important questions about the ability of costly punishment to promote the evolution of cooperation. Spiteful behavior including anti-social punishment is well documented aspects of human behavior (Saijo and Nakamura, 1995; Denant-Boemont et al., 2007; Dreber et al., 2008; Herrmann et al., 2008; Nikiforakis, 2008; Gächter and Herrmann, 2009; Wu et al., 2009), and should not be ignored. Here we show that in populations with local interaction and competition, anti-social and spiteful behaviors can be favored by evolution as they actually enhance an actor's relative reproductive success; and costly punishment no longer promotes the evolution of cooperation when anti-social punishment is allowed. Nonetheless, we observe cooperation throughout the natural world and human society. While many other models for the co-evolution of punishment and cooperation have been proposed, our results suggest that the effects of anti-social punishment need to be explored in these other contexts as well. If including anti-social punishment in other models gives results similar to what we have shown here, then mechanisms other than punishment could be primarily responsible for the evolution of cooperation, and punishment must have evolved secondarily for other reasons such as asserting dominance (Clutton-Brock and Parker, 1995a).

Acknowledgments

D.R. and J.A. acknowledge support from the John Templeton Foundation, the NSF/NIH joint program in mathematical biology (NIH Grant R01GM078986), the Bill and Melinda Gates Foundation (Grand Challenges Grant 37874), and J. Epstein. M.N. acknowledges Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (nos. 21770016, 20310086, 19046006, 19310097, 21247006) and the support from the Inamori Foundation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Derivation of Eq. (5) in the main text

Consider the game, (4). We study the stability condition of strategy X against Y. For that purpose let $\varepsilon(\ll 1)$ be the frequency of strategy Y in the population. The rest $1-\varepsilon$ is the frequency of X. A Y-strategist changes to strategy X at rate

$$\varepsilon\sum_{x=0}^{z}\binom{z}{x}(1-\varepsilon)^x\varepsilon^{z-x}d(xa_3+(z-x)a_4)\frac{x}{z}=\varepsilon d(za_3)+O(\varepsilon^2), \tag{A.1}$$

where $d(\cdot)$ is given by Eq. (3). In the sum above, x represents the number of X-players with whom a Y-player interacts. With probability $\binom{z}{x}(1-\varepsilon)^x\varepsilon^{z-x}$ the Y-player interacts with x many X-players and gains the payoff, $xa_3+(z-x)a_4$. When he dies, he changes to strategy X with probability x/z. On the other hand, an X-strategist changes to strategy Y at rate

$$(1-\varepsilon)\sum_{y=0}^{z}\binom{z}{y}(1-\varepsilon)^{z-y}\varepsilon^y d[(z-y)a_1+ya_2]\frac{y}{z}$$
$$=\varepsilon d((z-1)a_1+a_2)+O(\varepsilon^2). \tag{A.2}$$

There y represents the number of Y-players with whom an X-player interacts. With probability $\binom{z}{y}(1-\varepsilon)^{z-y}\varepsilon^y$ the X-player interacts with y many Y-players and gains the payoff, $(z-y)a_1+ya_2$. When he dies, he changes to strategy Y with probability y/z.

Comparing Eqs. (A.1) and (A.2), we obtain the stability condition of X against Y as $d(za_3)>d((z-1)a_1+a_2)$, or equivalently,

$$(z-1)a_1+a_2>za_3. \tag{A.3}$$

Appendix B. Conditions for DS strategy to win

An intuitive explanation of why the DS strategy wins in the z-mixed population model when β is relatively large is twofold: (i) there are no assortment effects in the z-mixed population model, such that one does not meets others with the same strategy more often than expected by global frequencies, and (ii) one learns from the same individuals one interacts with in the z-mixed population model, such that one's payoff relative to one's interaction partners matters.

To see that the property (i) is a necessary condition for DS to win, we compare the z-mixed population model (Section 3.1.2) with the lattice model (Section 3.2.2). Property (i) is present in the former but is absent in the latter, while property (ii) is present in both models. As we saw in the main text, DS wins in the z-mixed model but not in the lattice model, suggesting that property (i) is crucial for the propagation of DS strategy.

To see that (ii) is also a necessary condition for DS to win, we compare the z-mixed population model to a variant with 'global replacement'. In this variant, an updating player mimics the strategy of a player randomly chosen *from the entire population*, as opposed to a randomly selected interaction partner. Property (ii) is present in the z-mixed population model but is absent in the global replacement variant, whereas property (i) is present in both models. To compare the two models, we use agent based simulations with $z=4$, population size $N=2500$, $b=5$, $c=1$, $\alpha=1$, $\beta=10$ and viability updating parameter values $\gamma=1$ and $\theta=1$. We examine the outcome of 25 simulations lasting 125,000,000 generations, starting from an initial frequency of DS=0.79, all other strategies 0.03. Our calculations in the main text showed that DS is ESS in the z-mixed model when $\beta>z\alpha$. Consistent with this, we find that DS wins in the majority of cases in the z-mixed population agent based simulations. Conversely, we find that DN wins 80% of the time using the global replacement model. This result suggests that property (ii) is also crucial for the propagation of the DS strategy.

References

Alexander, R.D., 1987. The Biology of Moral Systems. Aldine de Gruyter, New York.
Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.
Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.
Axelrod, R., 1986. An evolutionary approach to norms. Am. Polit. Sci. Rev. 80, 1095–1111.
Boerlijst, M.C., Nowak, M.A., Sigmund, K., 1997. The logic of contrition. J. Theor. Biol. 185, 281–293.
Bowles, S., 2009. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? Science 324 1293–1298.
Bowles, S., Gintis, H., 2004. The evolution of strong reciprocity: cooperation in heterogeneous populations. J. Theor. Biol. 65, 17–28.
Boyd, R., Gintis, H., Bowles, S., Richerson, P.J., 2003. The evolution of altruistic punishment. Proc. Nat. Acad. Sci. USA 100, 3531–3535.
Boyd, R., Richerson, P.J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol. Sociobiol. 13, 171–195.
Brandt, H., Hauert, C., Sigmund, K., 2003. Punishment and reputation in spatial public goods games. Proc. R. Soc. London B 270, 1099–1104.
Brandt, H., Hauert, C., Sigmund, K., 2006. Punishing and abstaining for public goods. Proc. Nat. Acad. Sci. USA 103, 495–497.
Brandt, H., Sigmund, K., 2006. The good, the bad and the discriminator—errors in direct and indirect reciprocity. J. Theor. Biol. 239, 183–194.
Clutton-Brock, T.H., Parker, G.A., 1995a. Punishment in animal societies. Nature 373, 209–216.
Clutton-Brock, T.H., Parker, G.A., 1995b. Sexual coercion in animal societies. Anim. Behav. 49, 1345–1365.
Dal Bó, P., 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. Am. Econ. Rev. 95, 1591–1604.
Dal Bó, P., Frechette, G., 2009. The evolution of cooperation in infinitely repeated games: experimental evidence. Forthcoming in the Am. Econ. Rev. (See http://www.econ.brown.edu/fac/Pedro_Dal_Bo/).
Denant-Boemont, L., Masclet, D., Noussair, C., 2007. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. Econ. Theory 33, 1432–1479.
Dreber, A., Rand, D.G., Fudenberg, D., Nowak, M.A., 2008. Winners don't punish. Nature 452, 348–351.
El Mouden, C., Gardner, A., 2008. Nice natives and mean migrants: the evolution of dispersal-dependent social behaviour in viscous populations. J. Evol. Biol. 21, 1480–1491.
Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. Am. Econ. Rev. 90, 980–994.
Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140.
Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425, 785–791.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114, 817–868.
Foster, K.R., Wenseleers, T., Ratnieks, F.L.W., 2000. Spite in social insects. Trends Ecol. Evol. 15, 469–470.
Fowler, J.H., 2005. Altruistic punishment and the origin of cooperation. Proc. Nat. Acad. Sci. USA 102, 7047–7049.
Frank, S.A., 1998. Foundations of Social Evolution. Princeton University Press, Princeton.
Fu, F., Chen, X.J., Liu, L.H., Wang, L., 2007. Social dilemmas in an online social network: the structure and evolution of cooperation. Phys. Lett. A. 371 (1-2), 58–64.
Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. Am. Econ. Rev. 80, 274–279.
Fudenberg, D., Tirole, J., 1991. Game Theory. MIT Press, Cambridge, MA.
Fudenberg, D., Rand, D.G., Dreber, A., 2010. Slow to anger and fast to forgive: cooperation in an uncertain world. Available at SSRN: http://ssrn.com/abstract=1616396.
Gächter, S., Renner, E., Sefton, M., 2008. The long-run benefits of punishment. Science 322, 1510.
Gächter, S., Herrmann, B., 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. Philos. Trans. R. Soc. B 364, 791–806.
Gardner, A., West, S.A., 2004. Cooperation and punishment, especially in humans. Am. Nat. 164, 753–764.
Hamilton, W.D., 1964. The genetical evolution of social behaviour. J. Theor. Biol. 7, 1–52.
Hamilton, W.D., 1970. Selfish and spiteful behaviour in an evolutionary model. Nature 228, 1218–1220.
Hauert, C., Doebeli, M., 2004. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428, 643–646.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K., 2007. Via freedom to coercion: the emergence of costly punishment. Science 316, 1905–1907.
Hauser, M.D., 1992. Cost of deception: cheaters are punished in rhesus monkeys (*Macaca mulatta*). Proc. Nat. Acad. Sci. USA 89, 12137–12139.
Henrich, J., 2004. Cultural group selection, coevolutionary processes and large-scale cooperation. J. Econ. Behav. Organ. 53, 3–35.
Henrich, J., Boyd, R., 2001. Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. J. Theor. Biol. 208, 79–89.
Herrmann, B., Thoni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319, 1362–1367.
Irwin, A.J., Taylor, P.D., 2001. Evolution of altruism in stepping-stone populations with overlapping generations. Theor. Popul. Biol. 60, 315–325.
Janssen, M.A., Bushman, C., 2008. Evolution of cooperation and altruistic punishment when retaliation is possible. J. Theor. Biol. 254, 541–545.
Johnson, T., Dawes, C.T., Fowler, J.H., McElreath, R., Smirnov, O., 2009. The role of egalitarian motives in altruistic punishment. Econ. Lett. 102, 192–194.
Killingback, T., Doebeli, M., 1996. Spatial evolutionary game theory: Hawks and Doves revisited. Proc. R. Soc. B 263, 1135–1144.
Lehmann, L., Feldman, M., 2008. War and the evolution of belligerence and bravery. Proc. R. Soc. B 275, 2877–2885.
Lehmann, L., Keller, L., 2006. The evolution of cooperation and altruism—a general framework and a classification of models. J. Evol. Biol. 19, 1365–1376.
Lehmann, L., Bargum, K., Reuter, M., 2006. An evolutionary analysis of the relationship between spite and altruism. J. Evol. Biol. 19, 1507–1516.
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.
Lindgren, K., 1991. Evolutionary phenomena in simple dynamics. In: Langton, C. (Ed.), Artificial Life II. Addison-Wesley, Redwood City, CA, pp. 295–312.
Lindgren, K., Nordahl, M.G., 1994. Evolutionary dynamics of spatial games. Physica D 75, 292–309.
May, R.M., 1987. More evolution of cooperation. Nature 327, 15–17.
Molander, P., 1985. The optimal level of generosity in a selfish uncertain environment. J. Conflict Resolut. 29, 611–618.
Nakamaru, M., Dieckmann, U., 2009. Runaway selection for cooperation and strict-and-severe punishment. J. Theor. Biol. 257, 1–8.
Nakamaru, M., Matsuda, H., Iwasa, Y., 1997. The evolution of cooperation in a lattice structured population. J. Theor. Biol. 184, 65–81.
Nakamaru, M., Iwasa, Y., 2005. The evolution of altruism by costly punishment in lattice-structured populations: score-dependent viability versus score-dependent fertility. Evol. Ecol. Res. 7, 853–870.
Nakamaru, M., Iwasa, Y., 2006. The coevolution of altruism and punishment: role of the selfish punisher. J. Theor. Biol. 240, 475–488.
Nikiforakis, N., 2008. Punishment and counter-punishment in public goods games: can we still govern ourselves? J. Public Econ. 92 91–112.
Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.
Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359, 826–829.
Nowak, M.A., Sigmund, K., 1989. Game-dynamical aspects of the prisoner's dilemma. Appl. Math. Comput. 30, 191–213.
Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous populations. Nature 355, 250–253.
Nowak, M.A., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit for tat in prisoner's dilemma. Nature 364, 56–58.
Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437, 1291–1298.
Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs and social networks. Nature 441, 502–505.
Okasha, S., 2006. Evolution of the Levels of Selection. Oxford University Press, Oxford.
Oliver, P., 1980. Rewards and punishment as selective incentives for collective action: theoretical investigations. Am. J. Sociol. 85, 1356–1375.
Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: self-governance is possible. Am. Pol. Sci. Rev. 86, 404–417.
Page, K.M., Nowak, M.A., Sigmund, K., 2000. The spatial ultimatum game. Proc. R. Soc. B 267, 2177–2182.
Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432, 499–502.
Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A., 2009a. Positive interactions promote public cooperation. Science 325, 1272–1275.
Rand, D.G., Ohtsuki, H., Nowak, M.A., 2009b. Direct reciprocity with costly punishment: generous tit-for-tat prevails. J. Theor. Biol. 256, 45–57.
Ratnieks, F.L.W., Foster, K.R., Wenseleers, T., 2006. Conflict resolution in insect societies. Annu. Rev. Entomol. 51, 581–608.
Saijo, T., Nakamura, H., 1995. The "Spite" dilemma in voluntary contribution mechanism experiments. Journal of Conflict Resolution 39, 535–560.
Santos, F.C., Pacheco, J.M., Lenaerts, T., 2006. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. Proc. Nat. Acad. Sci. USA 103, 3490–3494.
Santos, F.C., Santos, M.D., Pacheco, J.M., 2008. Social diversity promotes the emergence of cooperation in public goods games. Nature 454, 213–216.
Sekiguchi, T., Nakamaru, M., 2009. Effect of the presence of empty sites on the evolution of cooperation by costly punishment in spatial games. J. Theor. Biol. 256, 297–304.
Shinada, M., Yamagishi, T., Ohmura, Y., 2004. False friends are worse than bitter enemies: "altruistic" punishment of in-group members. Evol. Hum. Behav. 25, 379–393.
Sigmund, K., 2007. Punish or perish? Retaliation and collaboration among humans. Trends Ecol. Evol. 22, 593–600.
Sigmund, K., Hauert, C., Nowak, M.A., 2001. Reward and punishment. Proc. Nat. Acad. Sci. USA 98, 10757–10762.
Szabo, G., Toke, C., 1998. Evolutionary prisoner's dilemma game on a square lattice. Phys. Rev. E 58, 69–73.
Taylor, P.D., 1992. Altruism in viscous populatons? An inclusive fitness model. Evol. Ecol. 6, 352–356.
Tibbetts, E.A., Dale, J., 2004. A socially enforced signal of quality in a paper wasp. Nature 432, 218–222.
Traulsen, A., Hauert, C., Brandt De Silva, H., Nowak, M.A., Sigmund, K., 2009. Exploration dynamics in evolutionary games. Proc. Nat. Acad. Sci. USA 106, 709–712.
Traulsen, A., Nowak, M.A., 2006. Evolution of cooperation by multilevel selection. Proc. Nat. Acad. Sci. USA 103, 10952–10955.
Trivers, R.L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57.
Wahl, L.M., Nowak, M.A., 1999. The continuous prisoner's dilemma: II. Linear reactive strategies with noise. J. Theor. Biol. 200, 323–338.
Wedekind, C., Milinski, M., 1996. Human cooperation in the simultaneous and the alternating prisoner's dilemma: pavlov verus generous tit-for-tat. Proc. Nat. Acad. Sci. USA 93, 2686–2689.
Wenseleers, T., Ratnieks, F.L.W., 2006. Enforced altruism in insect societies. Nature 444, 50.
West, S.A., Griffin, A.S., Gardner, A., 2007. Evolutionary explanations for cooperation. Curr. Biol. 17, R661–R672.
West, S.A., Gardner, A., 2010. Altruism, spite, and greenbeards. Science 327, 1341–1344.
Wu, J., Zhang, B., Zhou Z., He, Q., Zheng, X., Cressman, R., Tao, Y., 2009. Costly punishment does not always increase cooperation, Proc. Nat. Acad. Sci. USA, doi: 10.1073/pnas.0905918106.
Wilson, D.S., 1975. A theory of group selection. Proc. Nat. Acad. Sci. USA 72, 143–146.
Wilson, D.S., Pollock, G.B., Dugatkin, L.A., 1992. Can altruism evolve in purely viscous populations? Evol. Ecol. 6 331–341.
Wilson, D.S., Wilson, E.O., 2007. Rethinking the theoretical foundation of sociobiology. Q. Rev. Biol. 82, 327–348.
Wilson, E.O., 1975. Sociobiology: The New Synthesis. Harvard Press, Cambridge, Mass 697pp.
Yamagishi, T., 1986. The provision of a sanctioning system as a public good. J. Pers. Soc. Psychol. 51, 110–116.