

# Punishment does not promote cooperation under exploration dynamics when anti-social punishment is possible

Oliver P. Hauser<sup>1,2</sup>, Martin A. Nowak<sup>1,2,3</sup>, David G. Rand<sup>4,5,6,7\*</sup>

<sup>1</sup>Program for Evolutionary Dynamics, <sup>2</sup>Department of Organismic and Evolutionary Biology, <sup>3</sup>Department of Mathematics, Harvard University, Cambridge MA USA, <sup>4</sup>Department of Psychology, <sup>5</sup>Department of Economics, <sup>6</sup>Program in Cognitive Science, <sup>7</sup>School of Management, Yale University, New Haven CT USA

\*Corresponding author: david.rand@yale.edu

**It has been argued that punishment promotes the evolution of cooperation when mutation rates are high (i.e. when agents engage in ‘exploration dynamics’). Mutations maintain a steady supply of agents that punish free-riders, and thus free-riders are at a disadvantage. Recent experiments, however, have demonstrated that free-riders sometimes also pay to punish cooperators. Corresponding theoretical work has shown that often punishment no longer promotes cooperation when this ‘anti-social punishment’ is allowed. Here we study the effect of allowing all combinatorially possible punishment strategies in public goods games across a range of mutation rates. We find that for intermediate and high mutation rates, adding punishment does not promote cooperation in either compulsory or voluntary public goods games if anti-social punishment is possible: Mutations generate agents that punish cooperators just as frequently as agents that punish defectors, and these two effects cancel each other out. These results raise questions about the effectiveness of punishment for promoting cooperation when mutations are common, and highlight how decisions about which strategies to include in the strategy set can have profound effects on the resulting dynamics.**

Keywords: cooperation, anti-social punishment, mutation rates, evolutionary dynamics, exploration dynamics

## 1. Introduction

The evolution of cooperation is a central topic of interest across the natural and social sciences (Antal et al., 2009; Apicella et al., 2012; Axelrod, 1984; Capraro, 2013; Chudek and Henrich, 2011; Dal Bó, 2005; Dal Bó and Fréchette, 2011; Fudenberg and Maskin, 1990; Hauert and Doebeli, 2004; Hauert et al., 2002b; Helbing and Yu, 2009; Herrmann et al., 2008; Jacquet et al., 2011; Janssen et al., 2010; Levin, 2006; Milinski et al., 2002; Nowak and May, 1992; Nowak and Sigmund, 1992; Nowak and Sigmund, 1993; Nowak and Sigmund, 1998; Ostrom, 1990; Perc and Szolnoki, 2010; Peysakhovich and Rand, 2013; Rand et al., 2012; Rand et al., 2013; Rand et al., 2009b; Rapoport and Chammah, 1965; Rockenbach and Milinski, 2006; Seinen and Schram, 2006; Sigmund, 2010; Skyrms, 1996; Skyrms and Pemantle, 2000; Tarnita et al., 2009; Traulsen and Nowak, 2006; van Veelen et al., 2012; Wedekind and Milinski, 2000; Yoeli et al., 2013). Five mechanisms for the evolution of cooperation have been proposed: direct and indirect reciprocity, spatial selection, kin selection, and multi-level selection (Nowak, 2006; Rand and Nowak, 2013). Adding any of these interaction structures to a Prisoner's Dilemma can result in cooperation being favoured over defection. In recent years, the idea that individuals pay a personal cost to impose costs on others has gained increasing attention. Behavioural experiments have shown that individuals are willing to pay to punish others, and that this costly punishment often (although not always) results in an increase in cooperation (Almenberg et al., 2011; Dreber et al., 2008; Espín et al., 2012; Fehr and Gächter, 2000; Fehr and Gächter, 2002; Fehr and Fischbacher, 2004; Gächter et al., 2008; Gurerk et al., 2006; Herrmann et al., 2008; Janssen et al., 2010; Ostrom et al., 1992; Rand et al., 2009b; Rockenbach and Milinski, 2006; Sefton et al., 2007; Sutter et al., 2010; Ule et al., 2009; Yamagishi, 1986). Complimenting this empirical work is a body of literature using evolutionary game theory to explore the co-evolution of punishment and cooperation (Boyd and Richerson, 1992; Boyd et al., 2003; Fowler, 2005; Hauert et al., 2007; Helbing et al., 2010; Isakov and Rand, 2011; Nakamaru and Iwasa, 2005; Nakamaru and Iwasa, 2006; Ohtsuki et al., 2009; Rand et al., 2009a; Sigmund et al., 2001; Sigmund et al., 2010; Traulsen et al., 2009). These papers typically examine evolutionary outcomes when the ability to pay to punish defectors is added to one of the mechanisms for the evolution of cooperation: costly punishment is not itself a mechanism for the evolution of cooperation, but must be combined with reciprocity, spatial structure or group selection (Rand and Nowak, 2013).

The existence of a darker form of punishment, however, has begun to challenge the positive role of punishment suggested by much of this work. Numerous experimental studies have shown that some people also engage in ‘anti-social punishment’ directed at co-operators (Cinyabuguma et al., 2006; Ellingsen et al., 2012; Gächter and Herrmann, 2009; Gächter and Herrmann, 2011; Gächter et al., 2010; Herrmann et al., 2008; Nikiforakis, 2008; Rand and Nowak, 2011; Sylwester et al., (In press)). Because this anti-social punishment was seen as ‘illogical,’ it was excluded *a priori* from most previous theoretical models. Given the empirical evidence of anti-social punishment, however, it is important to re-evaluate previous models of the co-evolution of cooperation and punishment (García and Traulsen, 2012; Hilbe and Traulsen, 2012; Janssen and Bushman, 2008; Powers et al., 2012; Rand and Nowak, 2011; Rand et al., 2010). Furthermore, even in the absence of such empirical evidence, including all combinatorially possible strategies is essential so as to avoid biasing the model in favour of the behaviours one is trying to explain: whatever level of strategic complexity one chooses to consider, it is essential to include all possible strategies of that complexity.

Here we examine the effect of unbiasing the punishment strategy set when considering the co-evolution of cooperation and punishment under ‘exploration dynamics’ (i.e. high mutation rates). It has recently been suggested that cooperation can evolve via punishment when mutation rates are high (Traulsen et al., 2009). Frequent mutations serve to maintain all possible strategies at a high frequency in the population, regardless of fitness. In a model where the only possible punishment is targeted at defectors, therefore, mutations maintain a steady supply of punishers. As a result defectors fare poorly and are disfavoured.

We instead study a model in which punishment is not restricted to defectors only. When all forms of punishment are available, high mutation rates lead to a constant supply of individuals of all strategies, including those that punish cooperators. Thus cooperators are punished to the same extent as defectors, and punishment no longer promotes cooperation. We study compulsory and voluntary public goods games, as well as voluntary games where loners are exempt from punishment. In all cases, when mutations are common punishment does not promote the evolution of cooperation if anti-social punishment is not excluded.

The rest of the paper is structured as follows. In Section 2, we describe the payoff structure of the compulsory and voluntary public goods games. In Section 3, we describe the dynamics of the evolutionary process and the role of mutation rates. In Section 4, we provide results for

the compulsory game. In Section 5, we provide results for the voluntary game. In Section 6, we conclude.

## 2. Game payoffs

Let  $N$  denote the number of individuals in a population playing the Public Goods Game (PGG). The PGG is the multi-player version of the Prisoner's Dilemma (Hardin, 1968). Each player in a group of  $n$  players must decide whether or not to contribute a fixed amount  $c$  to the 'common good'. Contributions are multiplied by a factor  $r > 1$ , and evenly split by all group members, regardless of whether or not they contributed. Thus if  $x$  players choose to cooperate, each player receives  $rcx/n$  back from the common good. We refer to this game as the compulsory PGG because all players must participate in the PGG.

We also consider a voluntary PGG where participation in the game is not compulsory (Hauert et al., 2002a; Hauert et al., 2002b; Hauert et al., 2007). Players who choose to abstain from the game receive a constant loner's payoff  $\sigma$ , that is less than the  $(r-1)c$  payoff earned by each member of a group where everyone cooperates, but greater than the 0 payoff earned by each member of a group where everyone defects. If all but one player in a group are loners, the game cannot take place and everyone in the group receives the loner's payoff  $\sigma$ .

In both versions of the game, groups of size  $n$  are drawn from the population at random in each round to play a PGG. In the compulsory game, individuals choose between C (cooperate) or D (defect). In the voluntary game, a third option – L (loner) – is available. Following the PGG decision, each player is given the option to punish the others in the group: a player's punishment costs her a fee  $\gamma$  for every player she chooses to punish; a punished player incurs a cost  $\beta$  for every punishment she receives ( $\beta > \gamma$ ). Players can condition their punishment decisions on the PGG choice of each potential recipient of punishment.

We describe a strategy as  $W\text{-}XYZ$  (as in (Rand and Nowak, 2011)), where  $W \in \{C, D, L\}$  denotes the PGG decision,  $X \in \{N = "No", P = "Punish"\}$  denotes whether punishment is directed at cooperators,  $Y \in \{N, P\}$  whether it is directed at defectors, and  $Z \in \{N, P\}$  whether it is directed at loners. For example, C-NNP is a cooperator strategy that does not

punish cooperators and defectors but punishes loners, whereas D-PPP is a defector strategy that punishes all strategies (including others with the same strategy as itself). We now give the payoff that each strategy earns when playing against each other strategy for the compulsory and voluntary PGGs.

### *Compulsory Public Goods Game payoffs*

In the compulsory PGG, abstinence from the game (i.e., the loner strategy) is not an option. Thus there are, at most, 8 possible strategies:  $W=[C,D] \cdot X=[P,N] \cdot Y=[P,N]$  (to facilitate easier cross-comparison with the voluntary public goods game, we write  $Z=[N]$  for all strategies in the compulsory game as loners are non-existent and therefore not punishable). We define the payoff  $\pi_s$  of each strategy  $s$  as a function of the frequencies of cooperators  $X_C$ , defectors  $X_D$ , punishers of cooperation  $X_{pC}$  and punishers of defection  $X_{pD}$ :

$$\begin{aligned}\pi_{C-NNN} &= B(\mathbf{X}) - F(N, n) - X_{pC}\beta D(N, n) \\ \pi_{C-NPN} &= B(\mathbf{X}) - F(N, n) - (X_{pC}\beta + X_D\gamma)D(N, n) \\ \pi_{C-PNN} &= B(\mathbf{X}) - F(N, n) - ((X_{pC} - 1)\beta + (X_C - 1)\gamma)D(N, n) \\ \pi_{C-PPN} &= B(\mathbf{X}) - F(N, n) - ((X_{pC} - 1)\beta + (X_C - 1 + X_D)\gamma)D(N, n) \\ \pi_{D-NNN} &= B(\mathbf{X}) - X_{pD}\beta D(N, n) \\ \pi_{D-NPN} &= B(\mathbf{X}) - ((X_{pD} - 1)\beta + (X_D - 1)\gamma)D(N, n) \\ \pi_{D-PNN} &= B(\mathbf{X}) - (X_{pD}\beta + X_C\gamma)D(N, n) \\ \pi_{D-PPN} &= B(\mathbf{X}) - ((X_{pD} - 1)\beta + (X_C + X_D - 1)\gamma)D(N, n)\end{aligned}$$

where  $B(\mathbf{X}) = \frac{rcX_C}{N-1} \left(1 - \frac{1}{n}\right)$  is the payoff from the compulsory public goods game,  $F(N, n) = \left(1 - \frac{r(N-n)}{n(N-1)}\right)c$  is the effective cost of contributing to the compulsory public goods game, and  $D(N, n) = \frac{n-1}{N-1}$ .

For the compulsory game, we study three conditions. In the ‘No punishment’ condition, punishment is not an option for either cooperators or defectors: the only possible strategies are C-NNN and D-NNN. In the ‘Pro-social punishment’ condition, only punishment of

defectors by cooperators is allowed: the possible strategies are C-NNN, C-NPN, and D-NNN. This condition is equivalent to the compulsory game with punishment in (Traulsen et al., 2009). Finally, in the ‘Full punishment’ condition, all 8 strategies are available.

### *Voluntary Public Goods Game payoffs*

In the voluntary PGG, the option to abstain from the public goods game is made available. Loners who abstain from the public goods game receive a constant payoff  $\sigma$  that is independent from the contributions to the common good.

There are 24 possible W-XYZ strategies: [C,D,L] • [P,N] • [P,N] • [P,N]. Their payoffs are

$$\begin{aligned}
\pi_{C-NNN} &= B'(\mathbf{X}) - F'(X_L)c - X_{pC}\beta D(N, n) \\
\pi_{C-NNP} &= B'(\mathbf{X}) - F'(X_L)c - (X_{pC}\beta + X_L\gamma)D(N, n) \\
\pi_{C-NPN} &= B'(\mathbf{X}) - F'(X_L)c - (X_{pC}\beta + X_D\gamma)D(N, n) \\
\pi_{C-NPP} &= B'(\mathbf{X}) - F'(X_L)c - (X_{pC}\beta + (X_D + X_L)\gamma)D(N, n) \\
\pi_{C-PNN} &= B'(\mathbf{X}) - F'(X_L)c - ((X_C - 1)\gamma + (X_{pC} - 1)\beta)D(N, n) \\
\pi_{C-PNP} &= B'(\mathbf{X}) - F'(X_L)c - ((X_C + X_L - 1)\gamma + (X_{pC} - 1)\beta)D(N, n) \\
\pi_{C-PPN} &= B'(\mathbf{X}) - F'(X_L)c - ((X_C + X_D - 1)\gamma + (X_{pC} - 1)\beta)D(N, n) \\
\pi_{C-PPP} &= B'(\mathbf{X}) - F'(X_L)c - ((X_C + X_D + X_L - 1)\gamma + (X_{pC} - 1)\beta)D(N, n) \\
\pi_{D-NNN} &= B'(\mathbf{X}) - X_{pD}\beta D(N, n) \\
\pi_{D-NNP} &= B'(\mathbf{X}) - (X_L\gamma + X_{pD}\beta)D(N, n) \\
\pi_{D-NPN} &= B'(\mathbf{X}) - ((X_D - 1)\gamma + (X_{pD} - 1)\beta)D(N, n) \\
\pi_{D-NPP} &= B'(\mathbf{X}) - ((X_D + X_L - 1)\gamma + (X_{pD} - 1)\beta)D(N, n) \\
\pi_{D-PNN} &= B'(\mathbf{X}) - (X_{pD}\beta + X_C\gamma)D(N, n) \\
\pi_{D-PNP} &= B'(\mathbf{X}) - (X_{pD}\beta + (X_C + X_D)\gamma)D(N, n) \\
\pi_{D-PPN} &= B'(\mathbf{X}) - ((X_D + X_D - 1)\gamma + (X_{pD} - 1)\beta)D(N, n) \\
\pi_{D-PPP} &= B'(\mathbf{X}) - ((X_D + X_D + X_L - 1)\gamma + (X_{pD} - 1)\beta)D(N, n) \\
\pi_{L-NNN} &= \sigma - X_{pL}\beta D(N, n)
\end{aligned}$$

$$\begin{aligned}
\pi_{L-NNP} &= \sigma - \left( (X_L - 1)\gamma + (X_{pL} - 1)\beta \right) D(N, n) \\
\pi_{L-NPN} &= \sigma - (X_D\gamma + X_L\beta)D(N, n) \\
\pi_{L-NPP} &= \sigma - \left( (X_D + X_L - 1)\gamma + (X_{pL} - 1)\beta \right) D(N, n) \\
\pi_{L-PNN} &= \sigma - (X_C\gamma + X_L\beta)D(N, n) \\
\pi_{L-PNP} &= \sigma - \left( (X_C + X_L - 1)\gamma + (X_{pL} - 1)\beta \right) D(N, n) \\
\pi_{L-PPN} &= \sigma - \left( (X_C + X_D)\gamma + X_{pL}\beta \right) D(N, n) \\
\pi_{L-PPP} &= \sigma - \left( (X_C + X_D + X_L - 1)\gamma + (X_{pL} - 1)\beta \right) D(N, n)
\end{aligned}$$

where the benefit of the PGG is given by

$$B'(\mathbf{X}) = \frac{rcX_C}{N - X_L - 1} \left( 1 - \frac{N}{n(N - X_L)} \right) + \binom{X_L}{n-1} \left( \sigma + \frac{rcX_C(X_L - n + 1)}{n(N - X_L - 1)(N - X_L)} \right)$$

and the effective cost of contributing is given by

$$F'(X_L) = 1 - \frac{r}{n} \frac{N-n}{N-X_L-1} + \binom{X_L}{n-1} \left( \frac{r}{n} \frac{X_L+1}{N-X_L-1} + r \frac{N-X_L-2}{N-X_L-1} - 1 \right).$$

If  $X_L \geq (N - 1)$ , then there is never more than 1 loner in a group, the public goods game is never played, and  $B(\mathbf{X}) = \sigma$  and  $F(X_L) = 0$ .

For the voluntary game, we study four conditions. In the ‘No punishment’ condition, punishment is not an option for cooperators, defectors or loners: the possible strategies are C-NNN, D-NNN and L-NNN. In the ‘Pro-social punishment’ condition, only punishment of defectors by cooperators is allowed: the possible strategies are C-NNN, C-NPN, D-NNN and L-NNN. This condition is equivalent to the voluntary game with punishment in (Traulsen et al., 2009). In the ‘Full punishment’ condition, all 24 strategies are available. Finally, in the ‘Loners alone’ condition, loners are not allowed to punish or to be punished. Thus the following 9 strategies are possible: 4 cooperator strategies (C-NNN, C-NPN, C-PNN, C-PPN), 4 defector strategies (D-NNN, D-NPN, D-PNN, D-PPN) and 1 loner strategy (L-NNN). This condition is inspired by (García and Traulsen, 2012), who propose that cooperation can be sustained if punishment is made available only to players that participate in the PGG.

### 3. Evolutionary Dynamics

We study the transmission of strategies through an evolutionary process. This process may be genetic evolution or social learning. In either case, we assume that strategies with higher payoffs are more likely to survive and reproduce, while strategies with lower payoffs are less likely to do so. Mutations during reproduction lead to the introduction of novel strategies (selected uniformly at random). In the context of social learning, mutations may represent either confusion regarding the strategy of a player you are trying to copy, or experimentation with new strategies. Specifically, the term ‘exploration dynamics’ has been introduced to describe evolution under high mutation rates (Traulsen et al., 2009; Traulsen et al., 2010).

We study the evolutionary process as a frequency-dependent Moran process (Nowak et al., 2004) with exponential fitness (Traulsen et al., 2008). In each round, one individual  $i$  is picked for reproduction proportional to his/her fitness  $e^{\pi_i}$  (where  $\pi_i$  is the payoff of individual  $i$ ), while another individual  $j$  is chosen with uniform distribution to die, changing his/her strategy. With probability  $u$ , a mutation occurs and individual  $j$  picks a novel strategy. With probability  $1 - u$ , individual  $j$  adopts the strategy of individual  $i$ .

We study the evolutionary outcomes over a range of mutation rates using agent-based simulations. We also study the  $u \rightarrow 1$  ‘high mutation rate’ limit analytically (Antal et al., 2009). Here, selection contributes only slightly to the evolutionary dynamics, which are largely dominated by mutation. As a result, all possible strategies are on average almost equally abundant at the same time (unlike the low mutation, weak selection limit where strategies are equally abundant on average but there are never more than two strategies present at the same time). Thus, if there are  $M$  strategies, the frequency of strategy  $s$  is a perturbation from  $1/M$  which is a linear function of the payoff obtained when all strategies are equally abundant (Antal et al., 2009). Specifically, the relative perturbation  $\Delta_s$  from  $1/M$  is given by  $\pi_s^* - \frac{\sum_j x_j^*}{M}$ , where  $\pi_x^*$  is the expected payoff of strategy  $x$  in a population where all strategies have abundance  $1/M$ . Therefore the strategy that is most common at  $u \rightarrow 1$  is the strategy with the highest payoff when  $u = 1$  (i.e. when all strategies have frequency  $1/M$ ). Note that this result holds regardless of selection strength.

#### 4. Compulsory PGG Results

We begin with analytical calculations in the high mutation limit. We ask under what conditions cooperation can be favoured by natural selection. Selection favours cooperation in the high mutation limit if the expected payoff averaged over all cooperative strategies minus the expected payoff averaged over all possible strategies,  $\Delta_C$  is positive (i.e. if co-operators out-earn non-cooperators on average).

Using the No Punishment strategy set, we find

$$\Delta_C = -\frac{Fc}{2}.$$

Thus regardless of the payoff parameters, selection disfavours cooperation. Using the Pro-social Punishment strategy set, however,

$$\Delta_C = \frac{D(N,n)N(2\beta-\gamma)}{18} - \frac{Fc}{3}.$$

If punishment is sufficiently effective,

$$\beta > \frac{6Fc + D(N,n)Ny}{2D(N,n)N},$$

cooperation can be favoured. Using the Full Punishment strategy set, however, the result is identical to the No Punishment case:

$$\Delta_C = -\frac{Fc}{2}$$

Once again cooperation is disfavoured regardless of payoffs, and thus punishment does not promote cooperation.

We also explore dynamics outside of the high mutation limit using agent based simulations. We fix  $N = 100$ ,  $n = 5$ ,  $r = 3$ ,  $\gamma = 0.3$ , and  $\omega = \beta = c = 1$ . For each strategy set, we simulate  $10^7$  generations, and calculate the time averaged frequency of each strategy over the second half of the simulation (Figure 1a,c,e). For comparison purposes, we also numerically evaluate the perturbations in the high mutation limit by calculating the expected payoff for C and D strategies using each strategy set (Figure 1b,d,f). The simulations show qualitative agreement with the high mutation limit over a wide range of mutation rates: punishment does not promote cooperation when the Full Punishment strategy set is used. We also note that selection disfavours punishment of any kind (Figure 2): the aggregate frequency of strategies that punish defectors never rises above neutrally (0.5 in this case) using either the Prosocial

Punishment or Full Punishment strategy sets, nor does the aggregate frequency of strategies that punish cooperation in the Full Punishment case.

Furthermore, as shown in the analytical calculations, the difference in expected payoffs between cooperators and defectors is exactly the same in the No Punishment (Figure 1b) and Full Punishment (Figure 1f) strategy sets. Adding all punishment options has no effect on the relative standing of cooperation and defection, but does result in both types of strategies receiving dramatically lower absolute expected payoffs as a result of the large amount of punishing enacted.

## 5. Voluntary PGG Results

We again begin with analytical calculations in the high mutation limit, and compare  $\Delta_C$  across our four voluntary public goods game conditions. Using the No Punishment strategy set, we find

$$\Delta_C = \frac{B'(X)}{3} - \frac{\sigma}{3} - \frac{2Fc}{3}.$$

Thus even in the absence of punishment, cooperation can be favoured if the returns on cooperation are sufficiently high:

$$B'(X) > \sigma + 2Fc.$$

Using the Pro-social Punishment strategy set,

$$\Delta_C = \frac{B}{4} - \frac{\sigma}{4} - \frac{Fc}{2} + \frac{D(N,n)N(\beta-\gamma)}{16}.$$

Thus, if the effect of punishment is greater than the cost,  $\beta > \gamma$ , cooperation is favoured over a wider range of  $B'(X)$  values than in the No Punishment case: here punishment promotes cooperation (as in (Traulsen et al., 2009)). Using the Full Punishment strategy set, however, we again obtain a result identical to the No Punishment case:

$$\Delta_C = \frac{B'(X)}{3} - \frac{\sigma}{3} - \frac{2Fc}{3},$$

thus punishment does not promote cooperation when all strategies are possible, as in the compulsory game. Finally, we consider the “loners alone” strategy set, where we find

$$\Delta_C = \frac{B'(X)}{9} - \frac{\sigma}{9} - \frac{5Fc}{9} - \frac{4D(N,n)N(\beta+\gamma)}{81}.$$

As can be seen, punishment *inhibits* cooperation in this strategy set:  $\Delta C$  is decreasing in both  $\beta$  and  $\gamma$ .

Turning to agent based simulations, we use the same parameters as for the compulsory PGG with additional parameter of  $\sigma = 1$  (loner's payoff), and simulate over  $10^7$  generations. We calculate the time averaged frequencies of all strategies in the second half of the simulation (Figure 3). The results are similar to those of the compulsory game. Regardless of mutation rate, cooperation is never favoured by selection in the Full Punishment strategy set where anti-social punishment strategies are included. Similarly, cooperation is disfavoured in the "Loners Alone" strategy set as long as mutation rates are sufficiently large. We also find that punishment of all kinds is disfavoured when mutation rates are sufficiently large (Figure 4).

## 6. Discussion

We have shown that when an unbiased strategy set is used, punishment does not promote the evolution of cooperation when mutation rates are large ('exploration dynamics'). We do not make the assumption that punishment is only restricted to be used by cooperators and directed at defectors, as previously assumed in many models. When all forms of punishment are available, anti-social punishment towards cooperators is as common as traditional punishment of defectors in the high mutation limit. Thus anti-social punishment cancels out the positive effects that pro-social punishment may otherwise provide.

These results emphasize the importance of which strategies are (or are not) included in the strategy set. The choice of strategies is always important. But when agents are largely selecting strategies at random, then one's choice of which strategies to include has an especially profound effect on the evolutionary outcomes: even strategies which perform extremely poorly will be played frequently, and thus can have a substantial impact on the evolutionary outcomes. Thus it is critical to not inadvertently bias one's results by selectively omitting certain strategies (for example, strategies that are possible but seem 'illogical').

Our findings are consistent with previous work on anti-social punishment in the low mutation limit using the same payoff structures we studied (Rand and Nowak, 2011). The extension to higher mutation rates is of substantial interest because it has been suggested that such mutation rates are a form of 'exploration' and innovation that play an important role in human learning (Traulsen et al., 2009; Traulsen et al., 2010). Under these circumstances of high mutation, evolutionary dynamics may favour qualitatively different strategies compared

to when mutation is rare (Antal et al., 2009). We see this on our results when considering the voluntary public goods game in which loners are not allowed to punish or be punished. This concept of ‘leaving the loners alone’ is an effective mechanism for promoting cooperation in the low mutation limit (García and Traulsen, 2012). We show here, however, that when mutations are common punishment does not promote cooperation even when loners are left alone. On the contrary, punishment actually inhibits cooperation (by giving the non-punished loners an advantage over both co-operators and defectors).

In sum, we have shown that punishment does not promote cooperation when mutations are common unless a biased strategy set is used. These results add to a growing body of literature using evolutionary game theory that calls into question the positive role of peer punishment in the evolution of cooperation (Dreber and Rand, 2012; Janssen and Bushman, 2008; Powers et al., 2012; Rand and Nowak, 2011; Rand et al., 2009a; Rand et al., 2010).

### Acknowledgements

We thank Julián García and Arne Traulsen for helpful comments and suggestions. O.P.H. is grateful to the department of Organismic and Evolutionary Biology at Harvard for fellowship support. Funding from the John Templeton Foundation is gratefully acknowledged.

## References

- Almenberg, J., Dreber, A., Apicella, C. L., Rand, D. G., 2011. Third Party Reward and Punishment: Group Size, Efficiency and Public Goods. *Psychology and Punishment*. Nova Science Publishers, Hauppauge, NY.
- Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C. E., Nowak, M. A., 2009. Mutation-selection equilibrium in games with multiple strategies. *Journal of theoretical biology* 258, 614-622.
- Apicella, C. L., Marlowe, F. W., Fowler, J. H., Christakis, N. A., 2012. Social networks and cooperation in hunter-gatherers. *Nature* 481, 497-501.
- Axelrod, R., 1984. *The Evolution of Cooperation*. Basic Books, New York.
- Boyd, R., Richerson, P. J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology* 13, 171–195.
- Boyd, R., Gintis, H., Bowles, S., Richerson, P. J., 2003. The evolution of altruistic punishment. *Proc Natl Acad Sci USA* 100, 3531-5.
- Capraro, V., 2013. A Model of Human Cooperation in Social Dilemmas. *PLoS ONE*.
- Chudek, M., Henrich, J., 2011. Culture gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences* 15, 218-226.
- Cinyabuguma, M., Page, T., Puttermann, L., 2006. Can second-order punishment deter perverse punishment? *Experimental Economics* 9, 265-279.
- Dal Bó, P., 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review* 95, 1591–1604.
- Dal Bó, P., Fréchette, G. R., 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. *American Economic Review* 101, 411-429.
- Dreber, A., Rand, D. G., 2012. Retaliation and antisocial punishment are overlooked in many theoretical models as well as behavioral experiments. *Behavioral and brain sciences* 35, 24-24.
- Dreber, A., Rand, D. G., Fudenberg, D., Nowak, M. A., 2008. Winners don't punish. *Nature* 452, 348-51.
- Ellingsen, T., Herrmann, B., Nowak, M. A., Rand, D. G., Tarnita, C. E., 2012. Civic Capital in Two Cultures: The Nature of Cooperation in Romania and USA. Available at SSRN: <http://ssrn.com/abstract=2179575>.
- Espín, A. M., Brañas-Garza, P., Herrmann, B., Gamella, J. F., 2012. Patient and impatient punishers of free-riders. *Proceedings of the Royal Society B: Biological Sciences* 279, 4923-4928, doi:10.1098/rspb.2012.2043.
- Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. *American Economic Review* 90, 980-994.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature* 415, 137-40.
- Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. *Evolution and Human Behavior* 25, 63-87.
- Fowler, J. H., 2005. Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci U S A* 102, 7047-9, doi:0500938102 [pii] 10.1073/pnas.0500938102.
- Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. *American Economic Review* 80, 274-279
- Gächter, S., Herrmann, B., 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 791-806.

- Gächter, S., Herrmann, B., 2011. The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia. *European Economic Review* 55, 193-210, doi:10.1016/j.eurocorev.2010.04.003.
- Gächter, S., Renner, E., Sefton, M., 2008. The Long-Run Benefits of Punishment. *Science* 322, 1510.
- Gächter, S., Herrmann, B., Thöni, C., 2010. Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 2651-2661, doi:10.1098/rstb.2010.0135.
- García, J., Traulsen, A., 2012. Leaving the loners alone: Evolution of cooperation in the presence of antisocial punishment. *Journal of Theoretical Biology* 307, 168-173.
- Gurerk, O., Irlenbusch, B., Rockenbach, B., 2006. The competitive advantage of sanctioning institutions. *Science* 312, 108-11.
- Hardin, G., 1968. The Tragedy of the Commons. *Science* 162, 1243 - 1248.
- Hauert, C., Doebeli, M., 2004. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428, 643-646.
- Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002a. Replicator dynamics for optional public good games. *J Theor Biol* 218, 187-94.
- Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002b. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* 296, 1129-32.
- 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.
- Helbing, D., Yu, W., 2009. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences* 106, 3680-3685, doi:10.1073/pnas.0811503106.
- 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.
- Herrmann, B., Thoni, C., Gächter, S., 2008. Antisocial punishment across societies. *Science* 319, 1362-1367.
- Hilbe, C., Traulsen, A., 2012. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Sci. Rep.* 2, 458.
- Isakov, A., Rand, D. G., 2011. The Evolution of Coercive Institutional Punishment. *Dynamic Games and Applications*, DOI 10.1007/s13235-011-0020-9, doi:10.1007/s13235-011-0020-9.
- Jacquet, J., Hauert, C., Traulsen, A., Milinski, M., 2011. Shame and honour drive cooperation. *Biology letters* 7, 899-901.
- Janssen, M. A., Bushman, C., 2008. Evolution of cooperation and altruistic punishment when retaliation is possible. *Journal of theoretical biology* 254, 541-545.
- Janssen, M. A., Holahan, R., Lee, A., Ostrom, E., 2010. Lab Experiments for the Study of Social-Ecological Systems. *Science* 328, 613-617.
- Levin, S. A., 2006. Learning to live in a global commons: socioeconomic challenges for a sustainable environment. *Ecological Research* 21, 328-333, doi:10.1007/s11284-006-0162-1.
- Milinski, M., Semmann, D., Krambeck, H. J., 2002. Reputation helps solve the 'tragedy of the commons'. *Nature* 415, 424-6.
- 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. *Journal of theoretical biology* 240, 475-488.

- Nikiforakis, N., 2008. Punishment and Counter-punishment in Public Goods Games: Can we still govern ourselves? *Journal of Public Economics* 92, 91-112.
- Nowak, M. A., 2006. Five rules for the evolution of cooperation. *Science* 314, 1560-3.
- Nowak, M. A., May, R. M., 1992. Evolutionary games and spatial chaos. *Nature* 359, 826-829.
- 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 the Prisoner's Dilemma game. *Nature* 364, 56-58.
- Nowak, M. A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573-7, doi:10.1038/31225.
- Nowak, M. A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428, 646-650.
- Ohtsuki, H., Iwasa, Y., Nowak, M. A., 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature* 457, 79-82.
- Ostrom, E., 1990. Governing the commons: The evolution of institutions for collective action. Cambridge Univ Pr.
- Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: selfgovernance is possible. *The American Political Science Review* 86, 404-417.
- Perc, M., Szolnoki, A., 2010. Coevolutionary games--A mini review. *Biosystems* 99, 109-125.
- Peysakhovich, A., Rand, D. G., 2013. Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory. Available at SSRN: <http://ssrn.com/abstract=2294242>.
- Powers, S. T., Taylor, D. J., Bryson, J. J., 2012. Punishment can promote defection in group-structured populations. *Journal of Theoretical Biology* 311, 107-116.
- Rand, D. G., Nowak, M. A., 2011. The evolution of antisocial punishment in optional public goods games. *Nat Commun* 2, 434.
- Rand, D. G., Nowak, M. A., 2013. Human Cooperation. *Trends in Cognitive Sciences*, doi:10.1016/j.tics.2013.06.003.
- Rand, D. G., Ohtsuki, H., Nowak, M. A., 2009a. Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *J Theor Biol* 256, 45-57.
- Rand, D. G., Greene, J. D., Nowak, M. A., 2012. Spontaneous giving and calculated greed. *Nature* 489, 427-430.
- Rand, D. G., Armao IV, J. J., Nakamaru, M., Ohtsuki, H., 2010. Anti-social punishment can prevent the co-evolution of punishment and cooperation. *Journal of theoretical biology* 265, 624-632.
- Rand, D. G., Tarnita, C. E., Ohtsuki, H., Nowak, M. A., 2013. Evolution of fairness in the one-shot anonymous Ultimatum Game. *Proceedings of the National Academy of Sciences* 110, 2581-2586.
- Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M. A., 2009b. Positive Interactions Promote Public Cooperation. *Science* 325, 1272-1275.
- Rapoport, A., Chammah, A. M., 1965. *Prisoner's Dilema: A Study in Conflict and Cooperation*. University of Michigan Press.
- Rockenbach, B., Milinski, M., 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444, 718-23.
- Sefton, M., Schupp, R., Walker, J. M., 2007. The Effect of Rewards and Sanctions in Provision of Public Goods. *Economic Inquiry* 45, 671 - 690.
- Seinen, I., Schram, A., 2006. Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review* 50, 581-602.
- Sigmund, K., 2010. *The calculus of selfishness*. Princeton Univ Press, Princeton.

- Sigmund, K., Hauert, C., Nowak, M. A., 2001. Reward and punishment. *Proc Natl Acad Sci U S A* 98, 10757-62.
- Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. *Nature* 466, 861-863.
- Skyrms, B., 1996. Evolution of the social contract. Cambridge University Press.
- Skyrms, B., Pemantle, R., 2000. A dynamic model of social network formation. *Proceedings of the National Academy of Sciences* 97, 9340.
- Sutter, M., Haigner, S., Kocher, M. G., 2010. Choosing the Stick or the Carrot? Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies* 77, 1540-1566.
- Sylwester, K., Herrmann, B., Bryson, J. J., (In press). Homo homini lupus? Explaining antisocial punishment. *Journal of Neuroscience Psychology and Economics*.
- Tarnita, C. E., Antal, T., Ohtsuki, H., Nowak, M. A., 2009. Evolutionary dynamics in set structured populations. *Proceedings of the National Academy of Sciences* 106, 8601-8604, doi:10.1073/pnas.0903019106.
- Traulsen, A., Nowak, M. A., 2006. Evolution of cooperation by multilevel selection. *Proc Natl Acad Sci USA* 103, 10952-10955.
- Traulsen, A., Shores, N., Nowak, M., 2008. Analytical Results for Individual and Group Selection of Any Intensity. *Bulletin of Mathematical Biology* 70, 1410-1424.
- Traulsen, A., Hauert, C., De Silva, H., Nowak, M. A., Sigmund, K., 2009. Exploration dynamics in evolutionary games. *Proceedings of the National Academy of Sciences* 106, 709-712.
- Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H.-J., Milinski, M., 2010. Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences* 107, 2962-2966.
- Ule, A., Schram, A., Riedl, A., Cason, T. N., 2009. Indirect Punishment and Generosity Toward Strangers. *Science* 326, 1701-1704, doi:10.1126/science.1178883.
- van Veelen, M., García, J., Rand, D. G., Nowak, M. A., 2012. Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences* 109, 9929-9934.
- Wedekind, C., Milinski, M., 2000. Cooperation Through Image Scoring in Humans. *Science* 288, 850-852, doi:10.1126/science.288.5467.850.
- Yamagishi, T., 1986. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* 51, 110-116.
- Yoeli, E., Hoffman, M., Rand, D. G., Nowak, M. A., 2013. Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences* 110, 10424-10429.

## Figure Legends

**Figure 1. In the compulsory public goods game, defection dominates when antisocial punishment is allowed.** Shown is the frequency of strategies that cooperate (red) and defect (blue) from agent based simulations averaged over  $10^7$  generations using the No Punishment strategy set (a), the Prosocial Punishment strategy set (c) and the Full Punishment strategy set (e). Also shown are the expected payoffs in the high mutation limit averaged over all cooperator and defector strategies using the No Punishment strategy set (b), the Prosocial Punishment strategy set (d) and the Full Punishment strategy set (f). Parameters:  $N = 100, n = 5, r = 3, \gamma = 0.3, \omega = \beta = c = 1$  (except for b, d, and f:  $N = 96$ ).

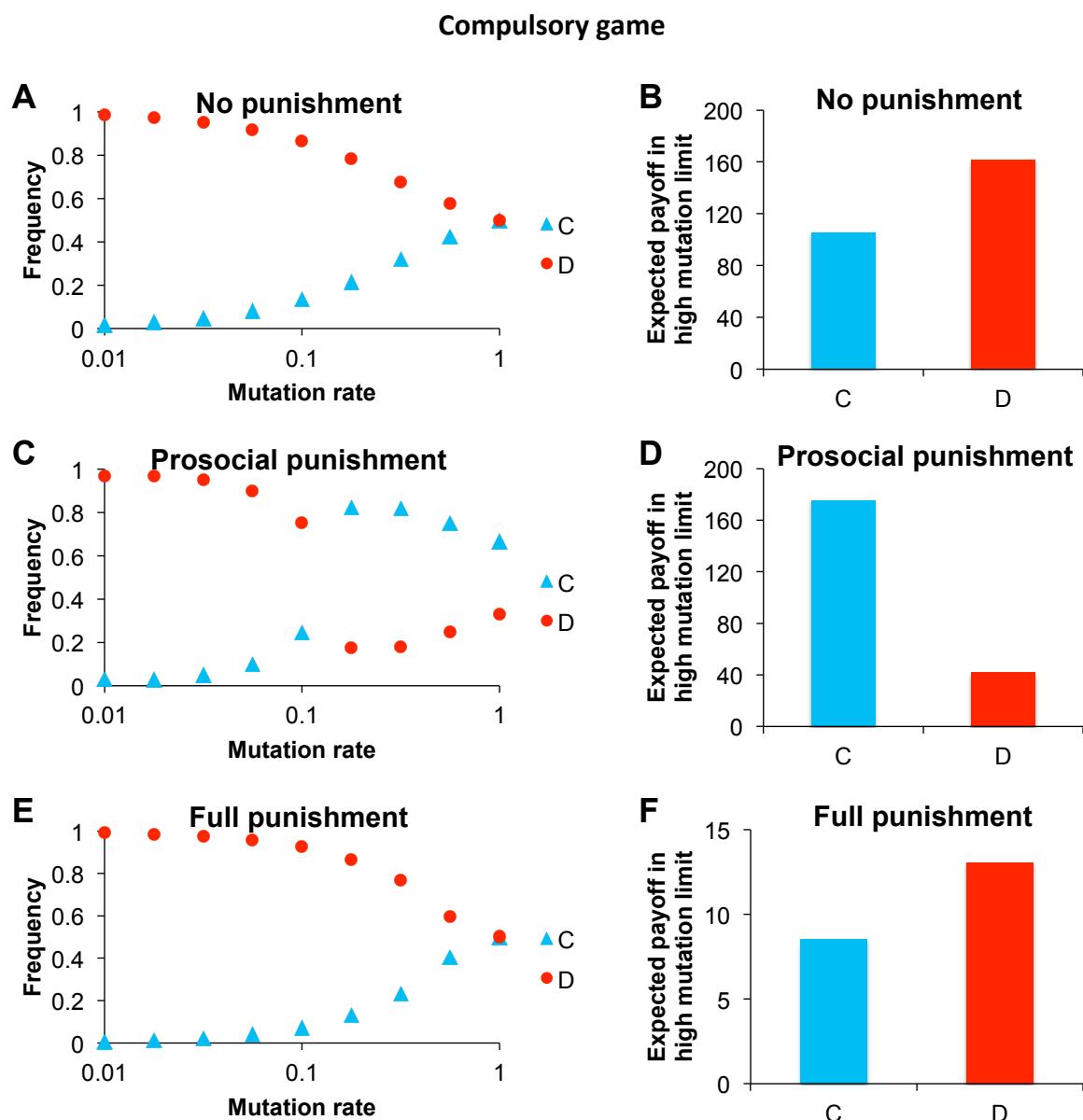
**Figure 2. Punishment is disfavoured by selection in the compulsory public goods game.** Shown is the frequency of strategies that punish cooperators (red) and punish defectors (blue) from agent based simulations averaged over  $10^7$  generations using the Prosocial Punishment strategy set (a) and the Full Punishment strategy set (b). Parameters:  $N = 100, n = 5, r = 3, \gamma = 0.3, \omega = \beta = c = 1$ .

**Figure 3. In the voluntary public goods game with antisocial punishment, cooperation is disfavoured when mutations are common.** Shown is the frequency of strategies that cooperate (red), defect (blue) and choose to be loners (yellow) from agent based simulations averaged over  $10^7$  generations using the No Punishment strategy set (a), the Prosocial Punishment strategy set (c), the Full Punishment strategy set (e) and the Loners Alone strategy set (g). Also shown are the expected payoffs in the high mutation limit averaged over all co-operator, defector and loner strategies using the No Punishment strategy set (b), the Prosocial Punishment strategy set (d), the Full Punishment strategy set (f) and the Loners Alone strategy set (h). Parameters:  $N = 100, n = 5, r = 3, \gamma = 0.3, \omega = \beta = \sigma = c = 1$  (except for b, d, f and g:  $N = 96$ ).

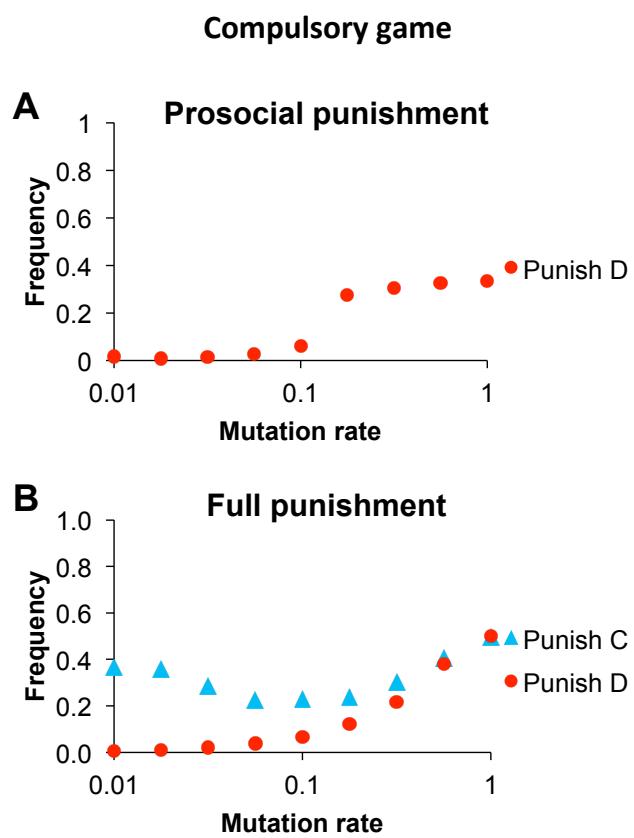
**Figure 4. Punishment is disfavoured by selection in the voluntary public goods game when mutation rates are large.** Shown is the frequency of strategies that punish cooperators (red), punish defectors (blue), and punish loners (yellow) from agent based simulations averaged over  $10^7$  generations using the Prosocial Punishment strategy set (a), the Full

Punishment strategy set (b), and the Loners Alone strategy set (c). Parameters:  $N = 100$ ,  $n = 5$ ,  $r = 3$ ,  $\gamma = 0.3$ ,  $\omega = \beta = c = 1$ .

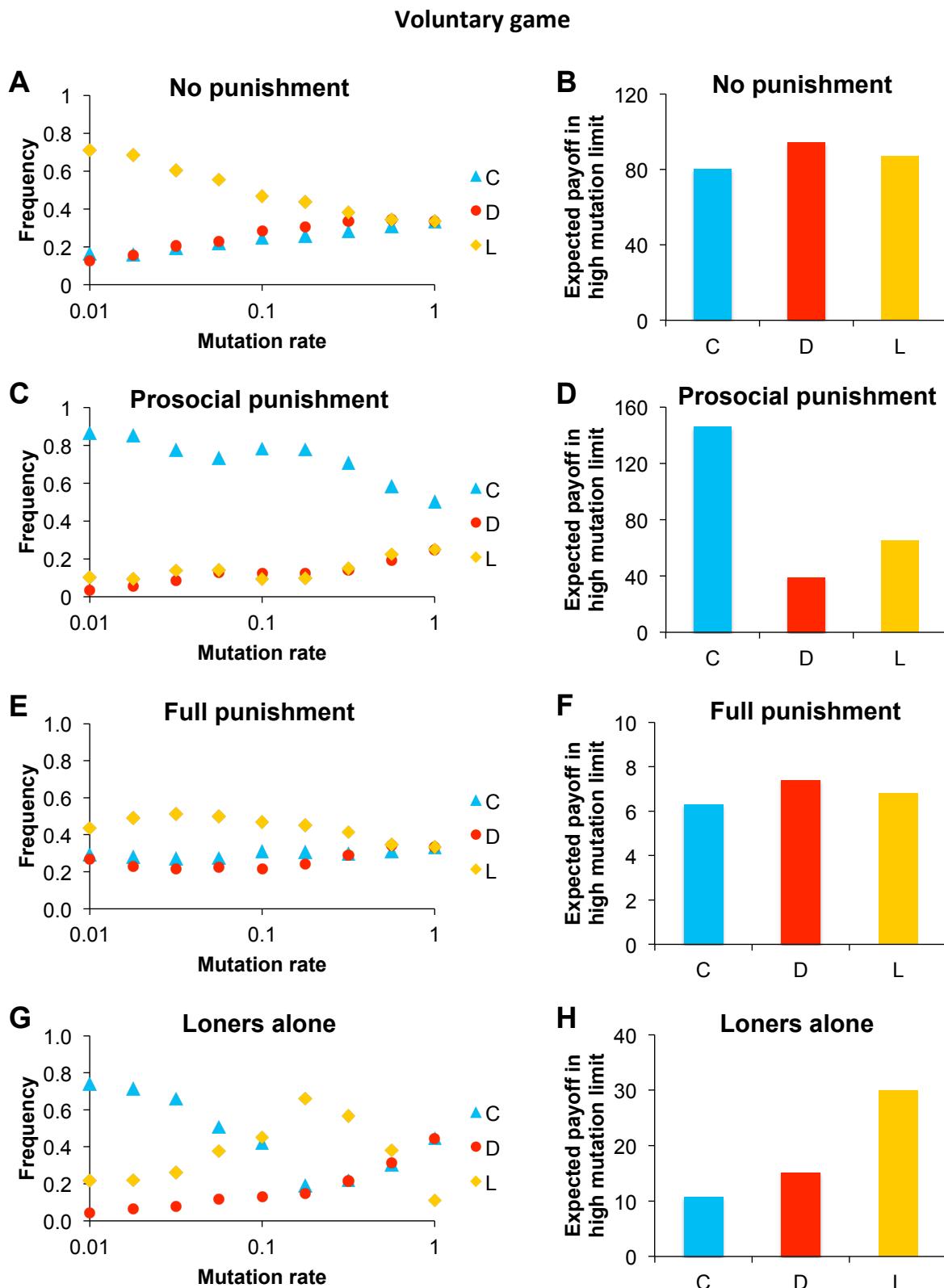
**Figure 1.**



**Figure 2.**



**Figure 3.**



**Figure 4.**

