

Public goods games with reward in finite populations

Peter A. I. Forsyth · Christoph Hauert

Received: 20 May 2010 / Revised: 24 August 2010
© Springer-Verlag 2010

Abstract Public goods games paraphrase the problem of cooperation in game theoretical terms. Cooperators contribute to a public good and thereby increase the welfare of others at a cost to themselves. Defectors consume the public good but do not pay its cost and therefore outperform cooperators. Hence, according to genetic or cultural evolution, defectors should be favored and the public good disappear – despite the fact that groups of cooperators are better off than groups of defectors. The maximization of short term individual profits causes the demise of the common resource to the detriment of all. This outcome can be averted by introducing incentives to cooperate. Negative incentives based on the punishment of defectors efficiently stabilize cooperation once established but cannot initiate cooperation. Here we consider the complementary case of positive incentives created by allowing individuals to reward those that contribute to the public good. The finite-population stochastic dynamics of the public goods game with reward demonstrate that reward initiates cooperation by providing an escape hatch out of states of mutual defection. However, in contrast to punishment, reward is unable to stabilize cooperation but, instead, gives rise to a persistent minority of cooperators.

Keywords Evolutionary game theory · Social dilemmas · Cooperation · Stochastic dynamics

Mathematics Subject Classification (2000) 91A22 · 82C31 · 60J10

P. A. I. Forsyth · C. Hauert (✉)
Department of Mathematics, University of British Columbia, 1984 Mathematics Road,
Vancouver, BC V6T 1Z2, Canada
e-mail: christoph.hauert@math.ubc.ca

Present Address:

P. A. I. Forsyth
Department of Economics, University of Toronto, 150 St. George Street,
Toronto, ON M5S 3G7, Canada

1 Introduction

Cooperation has the potential to improve the welfare of individuals but is a fragile accomplishment: when individuals are working together for mutual benefit, it is tempting to enjoy the fruits of the work of others while shirking one's own contribution. For example, groups of animals are well served if sentinels warn of approaching predators while others are busy foraging (Clutton-Brock et al. 1999) or if all members of the group participate in defending the young against predators (Hamilton 1971; Wilkinson and Shank 1977). However, in a group in which cooperation is the norm, individual animals may do better to shirk and forage in the safety of a watchful sentinel without returning the favor or to stand in the second line of a defense formation to reduce the risk of injury. Similarly, humans would be better off with sustainable management of natural resources such as air, drinking water, fossil fuels, fisheries, and the planet's climate, but uncertainties and the short term benefits of a few impede the implementation of effective environmental policies to the detriment of many (Milinski et al. 2008; Ostrom et al. 1999). Models of genetic and cultural evolution confirm that the opportunity to shirk undermines and often eliminates cooperation (Hardin 1968). Much work has been devoted to identifying and understanding mechanisms that are capable of promoting and stabilizing cooperation. In particular, the punishment of shirkers has proven to be an efficient way to maintain cooperation (Fehr and Gächter 2002; Sigmund 2007; Sigmund et al. 2001). However, punishment is costly and hence additional mechanisms are required to establish punishment (Brandt et al 2006; Colman 2006; Fowler 2005; Henrich et al. 2001).

Here we focus on the complementary approach and present a model in which individuals can attempt to promote cooperation by rewarding those that contribute to the public good. We demonstrate that reward is capable of promoting cooperation but is unable to stabilize it. Interestingly, significantly less effort has been dedicated to investigating positive incentives based on the rewarding of cooperators than has been devoted to the study of negative incentives based on the punishment of shirkers (De Silva and Sigmund 2009; Hauert 2010; Hilbe and Sigmund 2010; Oliver 1980; Sigmund et al. 2001). Game theoretical experiments indicate that, in general, a significant and reliable proportion of participants are willing to offer rewards for good behaviour (Andreoni et al. 2003; Fehr and Gächter 2000; Rand et al. 2009; Sefton et al. 2007).

2 Model

In the tradition of the extensive literature on punishment in public goods games, we consider interactions in two stages. Individuals first engage in a traditional public goods interaction (Kagel and Roth 1995) but the second stage now provides players with the opportunity to reward cooperating co-players as opposed to the opportunity to punish shirkers. In the first stage, each member in a group of N individuals may either cooperate and contribute to the common pool at some cost c or defect and not contribute anything. The sum of all contributions is multiplied by a factor r and evenly divided among all participants, irrespective of their contributions. With $1 < r < N$,

each invested dollar only provides a return of $r/N < 1$ to the investor and hence selfish individuals withhold their contributions in an attempt to free ride on the public good generated by others. Consequently, a group of selfish individuals foregoes the benefits of the public good and is worse off than a group of contributors, in which each individual ends up with $(r - 1)c$. This generates a conflict of interest between the individuals and the group that is called a social dilemma (Dawes 1980; Hauert et al. 2006).

In the second stage of the interaction, each individual may attempt to resolve the social dilemma by rewarding those that contributed to the public good. The provision of a reward entails costs γ but generates a benefit $\beta > \gamma$ to the rewarded individual. We model a situation in which an individual provides a service to others that is difficult or impossible for it to administer to itself, and therefore we exclude self-reward. An example of such a service is grooming behaviour (Stammbach and Kummer 1982). The two stage interaction results in four basic strategic types: the negligent cooperators—or second order free-riders— x , that cooperate but do not reward; the shirkers, y , that neither contribute nor reward; the pro-social individuals, v , that both cooperate and reward; as well as the risk averse individuals, w , that do not contribute but appreciate and reward the contributions of others. In order to ensure that rewarding is an attractive option, we require that (i) the benefits gained from one additional contributor to the public good exceed the costs of rewarding, $\gamma < rc/N$, and (ii) that, in a group of individuals that cooperate and reward, the accumulated net benefits of reward exceed the net costs of contributing to the public good, $(N - 1)(\beta - \gamma) > (1 - r/N)c$. Thus, whenever a group of at least $N - 1$ rewarders manages to entice an additional contribution to the public good, everybody is better off. Note that this includes the newly recruited contributor because of $(N - 1)\beta > (1 - r/N)c$ and hence the second condition eliminates the temptation to shirk cooperation among individuals that cooperate and reward. For $N = 2$, the public goods game with reward is similar to the trust game (Berg et al. 1995; De Silva and Sigmund 2009). The trust game, however, requires $\beta = \gamma$, which violates assumption (ii) above and hence creates a temptation to defect even in groups of rewarding cooperators.

Reward generates a second order dilemma because mutual reward increases the welfare of everyone but rewards are costly and hence selfish individuals should refrain from rewarding. But if no one rewards, the incentive to cooperate and contribute to the public good disappears. Thus, the dominant strategy is to refrain from rewarding and consequently also from contributing. The conclusion that neither cooperation nor reward can be sustained in social interactions is clearly at odds with our everyday observations. One way to resolve this problem is to note that the rationality assumptions used above can be easily challenged both for animal and for human populations. In nature, the strategies used in social interactions are rarely the product of rational calculations as the above analysis postulates, but are instead usually the outcome of an evolutionary process. In human society, rational calculation is present, but so is a tendency to imitate and adopt a strategy not because it is best, but because others succeed by using it. This tendency to imitate induces in societies a dynamic termed “cultural evolution”, which mimics genetic evolution. To capture the evolutionary dynamics in nature and human society, we introduce a stochastic evolutionary process, in which a finite population of individuals interact through public goods games with reward over

a number of generations. Those strategies that obtain the highest payoffs in any given generation tend to be adopted by a higher proportion of the population in the next generation—as Mailath puts it, the members of the population “are myopically groping toward maximizing behaviour” (Mailath 1998). In a cultural-evolutionary context, this means that successful individuals are more likely to be imitated; whereas in a genetic-evolutionary context, strategies are transmitted from parent to offspring, and successful individuals have a higher chance of producing progeny.

Another important component of social interactions, especially among humans, is reputation. In order to take this into account, we add to the model a small chance that an individual may learn whether its co-players reward contributions to the public good (Hauert 2010; Sigmund et al. 2001). This information could be obtained through gossip or observations of third party interactions. If a non-contributing individual realizes that all its co-players provide rewards, it opportunistically changes strategy to obtain these rewards. The condition $(N - 1)\beta > (1 - r/N)c$ ensures that it pays to do so.

2.1 Payoffs

Let $X, Y, V,$ and W be the number of individuals of the four strategic types $x, y, v,$ and w with $X + Y + V + W = M$, where M denotes the population size. Interaction groups of size N are randomly formed according to hypergeometric sampling. A non-contributor (types y, w) that is matched with i contributors (types x, v) and $N - i - 1$ other non-contributors receives, on average, a benefit B from the public goods interaction given by

$$B = \sum_{i=0}^{N-1} \frac{\binom{X+V}{i} \binom{Y+W-1}{N-1-i}}{\binom{M-1}{N-1}} \frac{rc}{N} i = \frac{rc}{N} (N - 1) \frac{X + V}{M - 1}. \tag{1}$$

In finite populations, the contributors’ benefit is slightly less because a contributor has one fewer contributor among its fellow members of the population than a non-contributor. Instead of discriminating between benefits to contributors and non-contributors we account for this difference in the effective costs of contribution. Hence, the average net benefit of contributors can be conveniently written as

$$\sum_{i=0}^{N-1} \frac{\binom{X+V-1}{i} \binom{Y+W}{N-1-i}}{\binom{M-1}{N-1}} \left(\frac{rc}{N} (i + 1) - c \right) = B - C, \tag{2}$$

with

$$C = \left(1 - \frac{r}{N} \left(1 - \frac{N - 1}{M - 1} \right) \right) c,$$

which takes into account that part of the investment returns to the investor as well as the effects of finite populations on the return to contributors. Similarly, in the second

stage of the interaction, a rewarding cooperator, v , receives on average a reward of

$$\sum_{i=0}^{N-1} \frac{\binom{V+W-1}{i} \binom{X+Y}{N-1-i}}{\binom{M-1}{N-1}} \beta i = \beta(N-1) \frac{V+W-1}{M-1} \tag{3}$$

from other rewarders (types v, t) and pays

$$-\gamma \sum_{i=0}^{N-1} \frac{\binom{X+V-1}{i} \binom{Y+W}{N-1-i}}{\binom{M-1}{N-1}} i = -\gamma(N-1) \frac{X+V-1}{M-1} \tag{4}$$

to reward fellow contributors (types x, v). An analogous calculation yields the corresponding reward for non-rewarding cooperators, x , as well as the costs of reward paid by rewarding non-contributors, w . This requires taking into account that in finite populations both types encounter one more rewarder or contributor, respectively, than the rewarding cooperators.

Let us now consider the effects of reputation. We model reputation by allowing individuals to temporarily adjust their contributions to the public good based on the reputation of their interaction partners. We assume that non-contributors are opportunistic: with probability ν a non-contributor in a group of rewarders learns that its interaction partners reward contributors and adjusts its behaviour by contributing to the public good. More precisely, a non-contributor adjusts its behaviour only if *all* partners offer rewards. The constraint that contributions are induced only if everybody rewards makes the following analysis more manageable, but we expect that there would be no qualitative differences if this threshold were lowered. For the same reasons, we assume that in every interaction group at most a single non-contributor may be persuaded to cooperate. This is a reasonable assumption provided that ν is small. With probability $1 - \nu$ the non-contributor has no information about whether it's co-players reward, and so does not contribute to the public good. This models a risk-averse strategy, in that a non-contributor contributes only if it knows doing so will improve its payoff. Alternatively, we could have implemented a propensity for non-contributors to decide to contribute based on inferences about their co-players' types. However, this would have considerably increased the model's complexity.

We have that with probability ν , a non-contributor (y or w) in a group with $N - 1$ rewarders (v or w) may learn that everybody rewards and take advantage of this information by deciding to cooperate. Rewarding cooperators benefit from the additional contribution to the public good but also incur the costs of rewarding one additional individual. If the non-contributor is of type y , the payoff to rewarders—weighted by the probability of this particular group composition—increases by

$$\nu \left(\frac{rc}{N} - \gamma \right) (N-1) \frac{Y}{M-1} \frac{\binom{V+W-1}{N-2}}{\binom{M-2}{N-2}} \tag{5}$$

and the payoff of the temporarily reformed shirker increases by

$$v \left(\beta(N - 1) - c + \frac{rc}{N} \right) \frac{\binom{V+W}{N-1}}{\binom{M-1}{N-1}}. \tag{6}$$

Similarly, if the non-contributor is of type w , the payoff to rewarders increases by

$$v \left(\frac{rc}{N} - \gamma \right) (N - 1) \frac{W}{M - 1} \frac{\binom{V+W-2}{N-2}}{\binom{M-2}{N-2}} \tag{7}$$

and the risk-averse individuals improve their lot by

$$v \left(\beta(N - 1) - c + \frac{rc}{N} \right) \frac{\binom{V+W-1}{N-1}}{\binom{M-1}{N-1}}. \tag{8}$$

Note the slightly different probabilities for the latter group configurations due to the finite size of the population. Finally, collecting all terms yields the expected payoffs for each strategic type:

$$P_x = B - C + \beta(N - 1) \frac{V + W}{M - 1} \tag{9a}$$

$$P_y = B + v \frac{\binom{V+W}{N-1}}{\binom{M-1}{N-1}} \left(\beta(N - 1) - c + \frac{rc}{N} \right) \tag{9b}$$

$$P_v = B - C + \beta(N - 1) \frac{V + W - 1}{M - 1} - \gamma(N - 1) \frac{X + V - 1}{M - 1} + v(N - 1) \left(\frac{rc}{N} - \gamma \right) \left(\frac{Y}{M - 1} \frac{\binom{V+W-1}{N-2}}{\binom{M-2}{N-2}} + \frac{W}{M - 1} \frac{\binom{V+W-2}{N-2}}{\binom{M-2}{N-2}} \right) \tag{9c}$$

$$P_w = B - \gamma(N - 1) \frac{X + V}{M - 1} + v \left(\frac{\binom{V+W-1}{N-1}}{\binom{M-1}{N-1}} \left(\beta(N - 1) - c + \frac{rc}{N} \right) + (N - 1) \left(\frac{rc}{N} - \gamma \right) \left(\frac{Y}{M - 1} \frac{\binom{V+W-1}{N-2}}{\binom{M-2}{N-2}} + \frac{W - 1}{M - 1} \frac{\binom{V+W-2}{N-2}}{\binom{M-2}{N-2}} \right) \right), \tag{9d}$$

Note that in the limit $M \rightarrow \infty$, $P_x + P_w = P_y + P_v$ holds, which implies that the dynamics admit a constant of motion (Hauert 2010; Hauert et al. 2004; Sigmund et al. 2001).

2.2 Dynamics

The state of the population is determined by the quadruple $\frac{1}{M}(X, Y, V, W)$, which takes values in $\{\frac{1}{M}(X, Y, V, W) \in \mathbb{R}^4 | (X, Y, V, W) \in \mathbb{Z}^4, (X, Y, V, W) \geq 0\}$,

$X + Y + V + W = M$ }, a subset of the simplex S_4 . Evolutionary change is modeled as a discrete-time stochastic process based on the Moran process (Moran 1962), which has recently been extended to populations of interacting individuals (Nowak et al. 2004). In each time step a focal individual is randomly selected from the population. With probability $1 - \theta$ it adopts the strategy of a model member i that is randomly chosen with a probability proportional to its fitness $f_i = 1 - \omega + \omega P_i$, a convex combination of the normalized baseline fitness of 1 and the payoff P_i from interactions with other members of the population. The parameter ω is called selection strength and determines the importance of fitness differences arising through different payoffs. Otherwise, with probability θ , the focal individual adopts a different strategy at random.

More specifically, if $\mathbf{Q} = (X, Y, V, W)$ denotes the state of the population, then the probability that, in any given time step, a focal individual of type j is chosen and changes to type $k \neq j$ through selection or exploration is given by

$$T_{j \rightarrow k}(\mathbf{Q}) = \frac{Q_j}{M} \left((1 - \theta) \frac{Q_k f_k}{\sum_i Q_i f_i} + \frac{\theta}{3} \right) \tag{10}$$

with $Q_x = X, Q_y = Y, Q_v = V, Q_w = W$. The first term inside the brackets describes success based selection, the driving force of genetic or cultural evolution, whereas the second term introduces a mutational process in which individuals randomly experiment with different strategies. Small θ correspond to rare genetic mutations, whereas large θ mimic exploratory behavior in cultural evolution (Traulsen et al. 2009).

The population dynamics are governed by a Master equation, where $\Omega^t(\mathbf{Q})$ denotes the probability that the population is in state \mathbf{Q} at time t :

$$\Omega^{t+1}(\mathbf{Q}) = \Omega^t(\mathbf{Q}) - T^{\text{out}}(\mathbf{Q})\Omega^t(\mathbf{Q}) + T^{\text{in}}(\mathbf{Q} + \delta)\Omega^t(\mathbf{Q} + \delta), \tag{11}$$

where $T^{\text{out}}(\mathbf{Q})$ summarizes the transition probabilities leading away from \mathbf{Q} , while $T^{\text{in}}(\mathbf{Q} + \delta)$ describes transitions from neighboring states $\mathbf{Q} + \delta$ into \mathbf{Q} (Traulsen et al. 2005, 2006, 2009).

The parameters θ and ω measure different forms of noise. ω measures a form of noise intrinsic to the selection process whereas θ measures an external source of noise arising from mutations or the propensity to explore other strategies. In the following we focus on the role of θ in order to compare cultural evolution (large θ) with genetic evolution (small θ). In order to reduce the effects of intrinsic noise we fix ω near its maximum. Note that for the payoff of cooperators $P_x < 0$ holds if shirkers abound (Y close to M) but fitness cannot be negative and hence $f_x = 0$ determines the maximum of ω .

3 Results

In general the solution to Eq. (11) is analytically inaccessible. However, the dynamics and the long-run average frequencies of the four strategic types can be explored through simulations as well as numerical and analytical approximations.

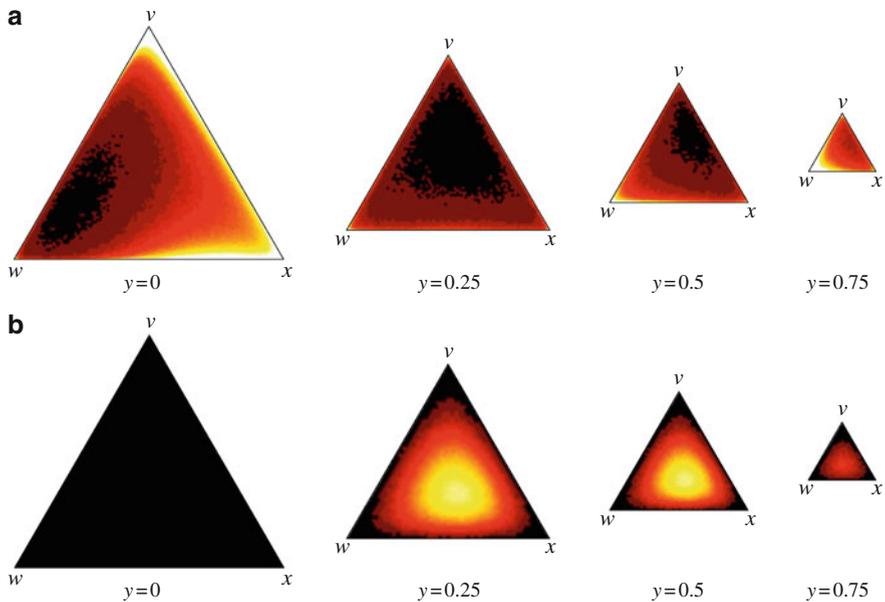
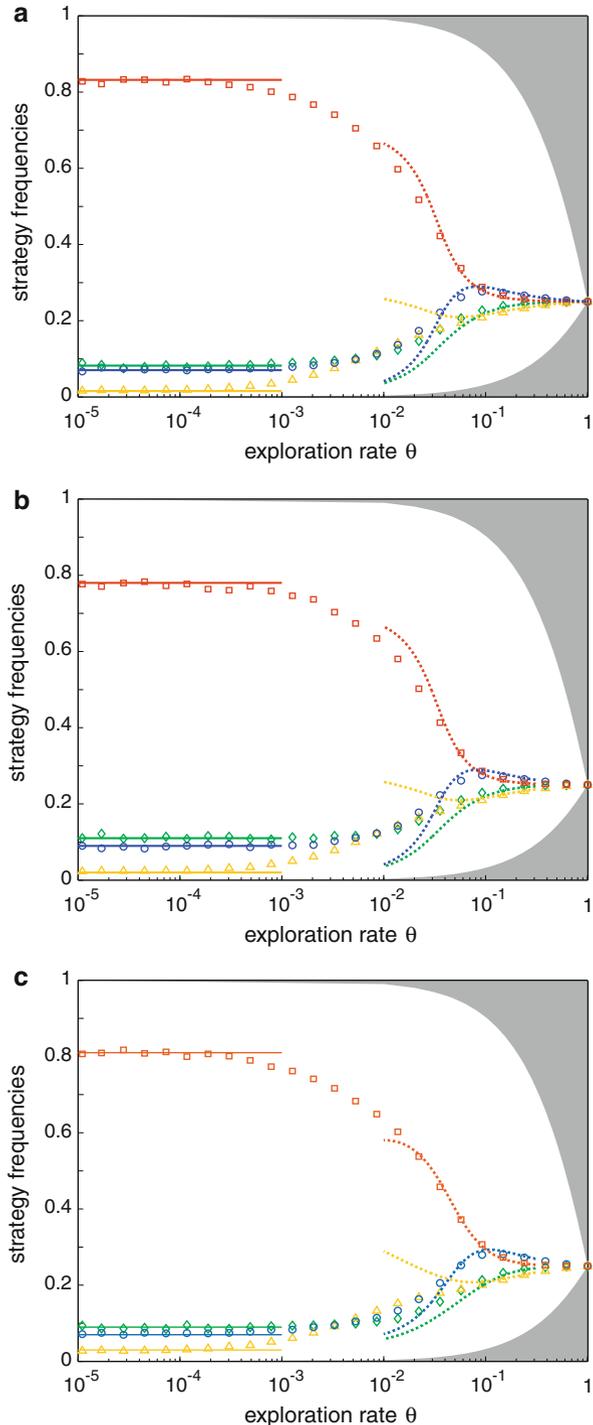


Fig. 1 Stochastic dynamics of public goods interactions with reward and reputation for different mutation or exploration rates θ . The simplex S_4 is sliced at different fractions of the shirkers y and the colors indicate the probability of the respective population composition (low black to white high). **a** For $\theta = 0.01$ the population spends most of the time along the boundary of S_4 whereas **b** for larger $\theta = 0.1$ the exploration process drives the population towards the interior. Parameters: $M = 100$, $N = 2$, $r = 1.2$, $c = 1$, $\gamma = 0.2$, $\beta = 1$, $v = 0.05$, $\omega = 0.4$; data collected over 10^8 generations. Results for $N = 5$ are essentially indistinguishable (not shown). Note that there are only finitely many states in our stochastic process. The distributions appear continuous because color has been interpolated between the states to make the figure easier to read (color figure online)

In our simulations, a population of size $M = 100$ is initialized with equal proportions of all four strategic types. The population is updated 10^8 times according to the dynamics described in Sect. 2.2, and the frequencies of the strategic types are averaged over all generations. Because of the stochastic nature of the evolutionary process and the large number of generations, the initial configuration is of negligible importance. Figure 1 shows the proportion of time the population spends in various states for $\theta = 0.01$ and $\theta = 0.1$. For $\theta = 0.01$ the population remains mostly close to the edges of S_4 , with one or two abundant types. However, for $\theta = 0.1$, the population spends most of its time in the interior of S_4 , with all strategies present. The symbols in Fig. 2 mark the long-run average frequencies based on the simulations.

The simulations are complemented by approximations of the dynamics in two limits: rare mutations ($\theta \rightarrow 0$), in which evolutionary changes are dominated by selection, and frequent explorations ($\theta \rightarrow 1$), in which selection is weak and evolutionary changes are dominated by random choices. For small mutation rates, the simulation results are in excellent agreement with the analytical rare mutation approximation, while for large mutation rates they agree well with the numerical frequent explorations approximation.

Fig. 2 Stochastic dynamics of public goods interactions with reward as a function of the exploration rate θ : **a** without reputation, $v = 0$; **b** with reputation, $v = 0.05$, for interactions in groups of $N = 5$, and **c** for pairwise interactions, $N = 2$ with reputation $v = 0.05$. The average frequencies of the four strategic types of negligent cooperators x (blue, open circle), shirkers y (red, open square), pro-social cooperators v (green, open diamond), and risk averse individuals w (yellow, open triangle) are shown together with analytical approximations in the rare mutation limit, $\theta \rightarrow 0$ (solid lines) as well as the numerical Fokker–Planck approximation in the limit of high exploration rates, $\theta \rightarrow 1$ (dashed lines). Because a fraction θ of the population always mutates, the minimum fraction of each strategy is $\theta/4$ and the grey shaded areas are inaccessible to the dynamics. Shirkers clearly dominate but more importantly they fail to eliminate cooperation altogether. Negligent and pro-social cooperators together consistently account for at least 15% and for $\theta > 0.1$ they even form the majority. Parameters: $M = 100, c = 1, \omega = 0.4$, **a, b** $r = 3, \gamma = 0.05, \beta = 0.25$; **c** $r = 1.2, \gamma = 0.2, \beta = 1$ (color figure online)



3.1 Rare mutations

Due to the stochastic nature of the evolutionary process, the population will eventually reach one of the four homogenous states, in which all individuals use the same strategy. At this point, the strategy that has been unanimously adopted is said to have fixated. If the mutation rate is zero ($\theta = 0$), the homogenous states are absorbing. For $\theta > 0$, a mutation will eventually cause the population to leave such a state. If the mutation rate is small, then the mutant strategy will either have fixated or disappeared before the next mutation arises (Antal and Scheuring 2006; Nowak et al. 2004; Ohtsuki et al. 2006). Thus, the population spends most of the time in these homogenous states and consists of at most two strategic types. The evolutionary dynamics reduce to a stochastic process along the edges of the simplex S_4 .

The direction of selection along each edge follows from pairwise comparisons of the payoffs. On the xy -edge $P_x - P_y = -C < 0$ holds, which simply reflects the fact that defectors dominate cooperators in the absence of opportunities to reward. Along the yw -edge, the two strategies are neutral for $v = 0$ because no one ever cooperates and hence no rewards are provided ($P_y = P_w$). However, for $v > 0$, w dominates provided that

$$P_w - P_y = v \frac{\binom{W-1}{N-2}}{\binom{M-2}{N-2}} (N-1) \left(\frac{rc}{N} - \gamma - \frac{(N-1)(\beta-\gamma) + (r-1)c - \frac{rc}{N} + \gamma}{M-1} \right) \quad (12)$$

is positive. For large M , the last term approaches zero and so the essential condition becomes $\frac{rc}{N} > \gamma$, i.e. whether the benefits from converting a non-contributor exceed the costs of rewarding one more individual. This is always satisfied for our initial assumptions. Also note that as long as $W < N - 1$, there are not enough rewarders to induce cooperation and the dynamics remain neutral until this threshold is reached. Along the wv -edge, v dominates provided that the total reward exceeds the effective costs of cooperation:

$$P_v - P_w = (1-v) \left((N-1) \left(\frac{\gamma}{M-1} + \beta \right) - C \right), \quad (13)$$

which is again satisfied by our initial assumptions. Next, along the vx -edge, x dominates because the negligent cooperators save the costs of rewarding $P_x - P_v = (N-1)(\gamma + \beta/(M-1)) > 0$. For $v > 0$, this generates a cyclic dominance of $x \rightarrow y \rightarrow w \rightarrow v \rightarrow x$. In the limit $M \rightarrow \infty$, this results in a heteroclinic cycle (Hauert 2010). The dynamics along the remaining two edges require a more careful analysis. Let us consider the vy -edge first:

$$P_v - P_y = -C + (N-1) \frac{V-1}{M-1} (\beta - \gamma) + v \frac{\binom{V-1}{N-2}}{\binom{M-2}{N-2}} \left[\frac{M-V}{M-1} (r-1)c - \left(1 + \frac{1}{M-1} \right) \left(\frac{rc}{N} - c + (N-1)\gamma \right) - (N-1) \frac{V}{M-1} (\beta - \gamma) \right]. \quad (14)$$

If V is small, Eq. (14) essentially reduces to $P_v - P_y \approx -C < 0$ and hence selection acts against the invasion of a rare v type. Conversely, if V is large (and v small), Eq. (14) approaches $P_v - P_y \approx -C + (N - 1)(\beta - \gamma)$, which is positive according to our initial assumptions and so selection also opposes the invasion of rare y types. Hence very few transitions along this edge are expected. Finally, along the wx -edge we have

$$\begin{aligned}
 P_w - P_x &= C - (N - 1) \left(\frac{W}{M - 1} \beta + \frac{M - W}{M - 1} \gamma \right) \\
 &\quad + v \binom{W-1}{N-1} \binom{M-1}{N-1} ((r - 1)c + (N - 1)(\beta - \gamma)). \tag{15}
 \end{aligned}$$

For large W , selection always favors the x type, but for small W , the w type may be favored if the total costs of rewarding are less than the costs of contributing to the public good, $(N - 1)\gamma < C$. If this is the case, both types can invade and mixtures of x and w may persist for extended times before fluctuations return the population to a homogenous state. If $(N - 1)\gamma > C$ then x dominates w .

In the limit of rare mutations, the population dynamics are well approximated by an embedded Markov chain based on the transition probabilities between the four homogenous states (Traulsen and Hauert 2009). These transitions solely depend on the dynamics along each edge, which are determined by the probability ϕ_{ij} that a single mutant of type i succeeds in invading and eventually replacing the resident population of type j (Nowak et al. 2004). The resulting transition matrix represents an irreducible finite Markov chain and hence the dynamics converge to a stationary distribution given by the eigenvector corresponding to the eigenvalue 1. The resulting estimates for average frequencies of the four types are in excellent agreement with the simulation data as marked by solid lines in Fig. 2. The stationary distribution and the dominant transitions for different parameter sets are illustrated in Fig. 3. In all cases the results are similar: the population is most often in the asocial state, but nevertheless it also spends a nontrivial amount of time in the cooperative states.

3.2 Frequent explorations

For high exploration rates, θ , and large population sizes, M , such that $\theta M \gg 1$, all strategies are always present and generate a fixed background for the stochastic selection process. In this limit the Master Eq. (11) can be approximated by a Fokker–Planck equation for the probability density $\rho(\mathbf{q})$ through a Kramers–Moyal expansion (Traulsen et al. 2005, 2006, 2009).

$$\dot{\rho}(\mathbf{q}) = - \sum_k \frac{\partial}{\partial q_k} \rho(\mathbf{q}) A_k(\mathbf{q}) + \frac{1}{2M} \sum_{j,k} \frac{\partial^2}{\partial x_k \partial x_j} \rho(\mathbf{q}) B_{jk}(\mathbf{q}), \tag{16}$$

with $\mathbf{q} = \mathbf{Q}/M$ and $j, k = x, y, v, w$. The deterministic component of the dynamics reflect fitness based selection and is given by the drift vector

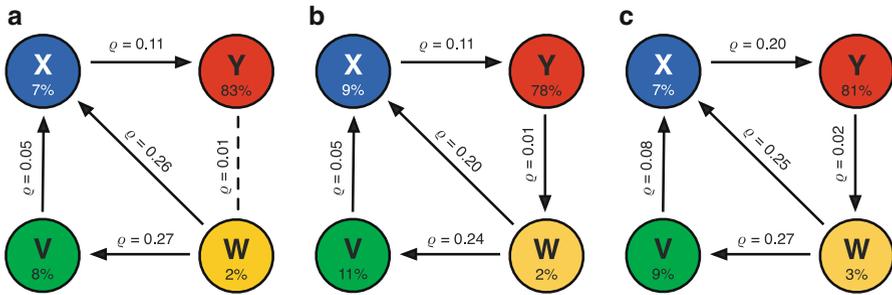


Fig. 3 Stochastic dynamics of public goods interactions with reward in the limit of rare mutations: **a** without reputation, $\nu = 0$; **b** with reputation, $\nu = 0.05$, for interactions in groups of $N = 5$ and **c** with reputation, $\nu = 0.05$, but for pairwise interactions, $N = 2$. Each circle marks one of the four homogeneous states with all negligent cooperators, X , shirkers, Y , pro-social cooperators, V , and risk averse individuals, W . The arrows indicate the dominant transitions ($>0.5\%$) and the dashed line in **a** marks the neutral edge, where no one ever cooperates. The selection strength ω is close to its maximum. In order to simplify comparisons for different group sizes, the parameters are chosen such that the maximum benefits and costs of reward, $(N - 1)\beta$ and $(N - 1)\gamma$, as well as the effective costs of cooperation, $(1 - r/N)c$, are equal in both settings. In spite of reward and reputation, the population clearly spends most of the time, $>78\%$, in the asocial shirker state where nobody cooperates and no rewards are provided. However, more importantly individuals keep cooperating 20% of the time even though they know the reputation of their partners only with a probability of 5%. Moreover, the fraction of cooperators barely changes for different group sizes. Parameters: $M = 100$, $c = 1$, $\omega = 0.4$, **a**, **b** $r = 3$, $\gamma = 0.05$, $\beta = 0.25$; **c** $r = 1.2$, $\gamma = 0.2$, $\beta = 1$

$$A_k(\mathbf{q}) = \sum_j (T_{j \rightarrow k} - T_{k \rightarrow j}). \tag{17}$$

The stochastic component of the dynamics is given by the diffusion matrix $B_{jk}(\mathbf{q})$, which characterizes the noise arising from random explorations and from the finite population size. For large M , effects of the noise term can be neglected and the stationary probability distribution has maxima close to states with $A_k(\mathbf{q}) = 0$. In the limit $\theta \rightarrow 1$ only a single maximum remains with $X = Y = V = W = M/4$. For θ close to 1, the deterministic drift term may admit several stable equilibria but the system spends the most time in the vicinity of the equilibrium with the largest basin of attraction. Hence the location of this equilibrium approximates the average state of the population. This approximation works well as long as there is a dominant basin of attraction and the noise facilitates easy transitions between the stable equilibria. The location of the dominant equilibrium can be numerically tracked for different θ . The dashed lines in Fig. 2 indicate the location of this equilibrium. Most notably, this approximation yields good estimates for $\theta > 0.05$ and confirms that for $\theta > 0.1$ contributors (cooperators and rewarders) are in the majority.

4 Discussion

The positive incentives generated by allowing individuals to reward those that contribute to the public good, combined with the stochastic dynamics of finite populations can indeed jump start cooperation. Note that in infinite populations, reward is potent only if

reputation is involved—otherwise asocial behavior rules and the positive incentives are disregarded (Hauert 2010; Sigmund et al. 2001). In contrast, reputation is no longer crucial in finite populations because stochastic fluctuations help to escape states of mutual defection and hence facilitate cooperation. However, even though reward can entice cooperation, it is incapable of maintaining it. Rewarding cooperators are prone to exploitation by second-order free-riders, who contribute to the public good but enjoy the rewards offered by their co-players without bothering to return the costly favors. In turn, this leads to the demise of cooperation and the successful return of asocial shirkers. Nevertheless, on average, a significant cooperating minority persists and this may serve as an important first step in promoting cooperation. This seems particularly relevant in connection with the complementary approach based on the punishment of those who fail to contribute to the public good. Once it is established, punishment is efficient because the mere threat of punishment may maintain cooperation at essentially no cost. However, it remains unclear how punishment behavior could gain a foothold in a population because, when rare, punishers have to punish left and right and hence fare poorly (Colman 2006).

In addition, punishment is prone to second-order free-riders. In the absence of reputation, contributors that are unwilling to bear the costs of punishment can undermine the efforts of punishers because they remain unnoticed. Nevertheless, they gradually lower the population's ability to resist the invasion of defectors (Brandt et al 2006; Hauert et al. 2007). Reputation helps to stabilize punishment (and cooperation) but makes establishing it even harder (Hauert et al. 2004; Sigmund et al. 2001). This further emphasizes the importance of mechanisms capable of overcoming the economic stalemate of mutual defection. Allowing for voluntary participation in joint enterprises provides such an escape hatch (Hauert et al. 2002) but introducing positive incentives through reward appears to be an interesting and relevant alternative.

The three key quantities determining the dynamics of public goods interactions with reward are (i) the effective costs of cooperation $(1 - r/N)c$, (ii) the peak costs of rewarding, $(N - 1)\gamma$, as well as (iii) the maximum reward, $(N - 1)\beta$. Keeping these quantities constant ensures that rewarding remains an attractive option, i.e. $\gamma < rc/N$ and $(N - 1)(\beta - \gamma) > (1 - r/N)c$, and essentially eliminates effects of different interaction group sizes, N . In fact, the average frequency of cooperators barely changes for different N (cf. Fig. 3b, c; Fig. 2b, c). This is particularly important because establishing cooperation in larger groups is particularly challenging (Boyd and Richerson 1988; Molander 1992). Yet, the effectiveness of positive incentives seems to be little affected by the group size.

The condition that the above quantities remain constant requires that the return on investments in the public good, r , scales with the group size, N . This holds whenever joint enterprises are subject to economies of scale such as in the hunting of large game or the construction of infrastructure (Alvard and Nolin 2002; Sugden 1986). In addition, the condition requires that the maximum costs and benefits of rewarding remain constant, but this actually becomes easier to satisfy as N increases because each individual's share in the total reward decreases.

Note that in the trust game ($N = 2, \gamma = \beta$) the condition $(N - 1)(\beta - \gamma) > (1 - r/N)c$ is violated and the asocial y -type can readily invade a population of pro-social v -types (see Eq. (14)). Consequently the public good cannot be maintained and

defectors dominate. In fact, the persistent minority of contributors in public goods games with reward crucially depends on the leverage provided by $\beta > \gamma$ and the fact that at least among prosocial types there is no incentive to withhold contributions—although the incentive to withhold rewards persists.

In the deterministic limit of small mutation rates and increasing population sizes reputation becomes increasingly important to overcome asocial states dominated by non-rewarding shirkers (Hauert 2010). Intuitively this means that the power of reward hinges on everybody becoming aware of its availability. However, when the dynamics are stochastic, individuals can figure this out by chance and hence reputation has only a small effect. In fact, at high exploration rates contributors (types x and v) dominate the population even in the absence of reputation. High exploration rates ensure a continuous supply of rewarding individuals, which benefits contributors but enables second-order free-riders to gain the upper hand because they evade the costs of rewarding (see Fig. 2). A related dominance of second-order free-riders at high exploration rates was observed for punishment (Traulsen et al. 2009): a constant background of punishers keeps the shirkers at bay and allows the non-punishing cooperators to thrive.

The concept of rewarding those that contribute to the public good is reminiscent of upstream or generalized reciprocity (Nowak and Roch 2006; Pfeiffer et al. 2005) in that acts of cooperation inspire more cooperation. This seems to be an important component of at least human behavior. Nevertheless, the experience of everyday life indicates that punishment is the more popular method of encouraging good behaviour, and indeed this may be because continually handing out rewards to all those who do not break the rules is tedious and costly, but, assuming that most people behave well, punishing those who deviate is cheap.

Acknowledgments C.H. acknowledges financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

- Alvard MS, Nolin DA (2002) Rousseau's whale hunt? Coordination among big-game hunters. *Curr Anthropol* 43(4):533–559
- Andreoni J, Harbaugh W, Vesterlund L (2003) The carrot or the stick: rewards, punishment, and cooperation. *Am Econ Rev* 93(3):893–902
- Antal T, Scheuring I (2006) Fixation of strategies for an evolutionary game in finite populations. *Bull Math Biol* 68:1923–1944
- Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. *Games Econ Behav* 10:122–142
- Boyd R, Richerson PJ (1988) The evolution of reciprocity in sizeable groups. *J Theor Biol* 132:337–356
- Brandt H, Hauert C, Sigmund K (2006) Punishing and abstaining for public goods. *Proc Natl Acad Sci USA* 103(2):495–497
- Clutton-Brock TH, O'Riain MJ, Brotherton PNM, Gaynor D, Kansky R, Griffin AS, Manser M (1999) Selfish sentinels in cooperative mammals. *Science* 284:1640–1644
- Colman A (2006) The puzzle of cooperation. *Nature* 440:744–745
- Dawes RM (1980) Social dilemmas. *Annu Rev Psychol* 31:169–193
- De Silva H, Sigmund K (2009) In games, groups and the global good. In: Levin SA (ed) *Game theory*. Springer, Berlin pp 85–114
- Fehr E, Gächter S (2000) Fairness and retaliation: the economics of reciprocity. *J Econ Perspect* 14(3):159–181
- Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415:137–140

- Fowler JH (2005) Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci USA* 102(19):7047–7049
- Hamilton WD (1971) The geometry of the selfish herd. *J Theor Biol* 31:295–311
- Hardin G (1968) The tragedy of the commons. *Science* 162:1243–1248
- Hauert C (2010) Replicator dynamics of reward & reputation in public goods games. *J Theor Biol* 267:22–28
- Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Volunteering as red queen mechanism for cooperation in public goods games. *Science* 296:1129–1132
- Hauert C, Haiden N, Sigmund K (2004) The dynamics of public goods. *DCDS-B* 4(3):575–587
- Hauert C, Michor F, Nowak M, Doebeli M (2006) Synergy and discounting of cooperation in social dilemmas. *J Theor Biol* 239:195–202
- Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K (2007) Via freedom to coercion: the emergence of costly punishment. *Science* 316:1905–1907
- Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R (2001) Cooperation, reciprocity and punishment in fifteen small scale societies. *Am Econ Rev* 91:73–78
- Hilbe C, Sigmund K (2010) Incentives and opportunism: from the carrot to the stick. *Proc R Soc Lond B* (online)
- Kagel JH, Roth AE (eds) (1995) *The handbook of experimental economics*. Princeton University Press, Princeton
- Mailath GJ (1998) Do people play nash equilibrium? Lessons from evolutionary game theory. *J Econ Lit* 36:1347–1374
- Milinski M, Sommerfeld RD, Krambeck H-J, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc Natl Acad Sci USA* 105(7):2291–2294
- Molander P (1992) The prevalence of free riding. *J Confl Resolut* 36:756–771
- Moran PAP (1962) *The statistical processes of evolutionary theory*. Clarendon, Oxford
- Nowak MA, Roch S (2006) Upstream reciprocity and the evolution of gratitude. *Proc R Soc Lond B* 274:605–609
- Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428:646–650
- Ohtsuki H, Hauert C, Lieberman E, Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441:502–505
- Oliver P (1980) Rewards and punishments as selective incentives for collective action: theoretical investigations. *Am J Soc* 85(6):1356–1375
- Ostrom E, Burger J, Field CB, Norgaard RB, Policansky D (1999) Revisiting the commons: local lessons, global challenges. *Science* 284:278–282
- Pfeiffer T, Rutte C, Killingback T, Taborsky M, Bonhoeffer S (2005) Evolution of cooperation by generalized reciprocity. *Proc R Soc Lond B* 272:1115–1120
- Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA (2009) Positive interactions promote public cooperation. *Science* 325:1272–1275
- Sefton M, Schupp R, Walker JM (2007) The effect of rewards and sanctions in provision of public goods. *Econ Inq* 45(4):671–690
- Sigmund K (2007) Punish or perish? Retaliation and collaboration among humans. *TREE* 22(11):593–600
- Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. *Proc Natl Acad Sci USA* 98:10757–10762
- Stammback E, Kummer H (1982) Individual contributions to a dyadic interaction: an analysis of baboon grooming. *Anim Behav* 30(4):964–971
- Sugden R (1986) *The economics of rights, co-operation and welfare*. Blackwell, Oxford
- Traulsen A, Claussen JC, Hauert C (2005) Coevolutionary dynamics: from finite to infinite populations. *Phys Rev Lett* 95:238701
- Traulsen A, Hauert C (2009) In: Schuster HG (ed) *Reviews of nonlinear dynamics and complexity*, vol II. Wiley-VCH, Weinheim
- Traulsen A, Hauert C, Silva HD, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. *Proc Natl Acad Sci USA* 106:709–712
- Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E* 74:011909
- Wilkinson GS, Shank CC (1977) Rutting-fight mortality among musk oxen on banks island, Northwest Territories, Canada. *Anim Behav* 24:756–758