

# Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games

M. ARCHETTI

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

## Keywords:

brinkmanship;  
conflict;  
cooperation;  
public good;  
social dilemma;  
social evolution;  
strategy;  
volunteer's dilemma.

## Abstract

Conflict and cooperation for the exploitation of public goods are usually modelled as an N-person prisoner's dilemma. Many social dilemmas, however, would be described more properly as a volunteer's dilemma, in which a certain number of individuals are necessary to produce a public good. If volunteering is costly, but so is failure to produce the public good, cheaters can invade and form a stable mixed equilibrium with cooperators. The dilemma is that the benefit for the group decreases with group size because the larger the group is, the less likely it is that someone volunteers. This problem persists even in the presence of a high degree of relatedness between group members. This model provides precise, testable predictions for the stability of cooperation. It also suggests a counterintuitive but practical solution for this kind of social dilemmas: increasing the damage resulting from the failure to produce the public good increases the probability that the public good is actually produced. Adopting a strategy that entails a deliberate risk (brinkmanship), therefore, can lead to a benefit for the society without being detrimental for the individual.

## Introduction

### Social dilemmas

Social dilemmas are situations in which the optimal decision of an individual contrasts with the optimal decision for the group. In game theory, this usually means games in which a dominant strategy leads to a Pareto inefficient equilibrium (Hardin, 1968; Dawes, 1980); the prisoner's dilemma (PD) (Tucker, 1950; Luce & Raiffa, 1957) is probably the most famous example. Solutions to these social dilemmas require repeated interactions, which allow reciprocation, punishment and reputation effects (Axelrod & Hamilton, 1981; Nowak, 2006). Situations of conflict for the exploitation of common resources (public goods games) are usually modelled as an N-person version of the PD. Individuals can be cooperators or defectors; cooperators pay a cost for contributing to the public good, whereas defectors refrain from doing so; after all individuals are given the chance to contribute to the

public good, the accumulated contribution is multiplied by an enhancement factor, and the total amount equally shared among all individuals (cooperators and defectors). As in the PD, because an individual can always exploit the benefits of living in a group without contributing to the costs, defection is the dominant strategy.

Although the PD and its N-person version have been the most popular games in the study of the evolution of cooperation, there are social dilemmas for which a different game would be a more appropriate model. For example, when viruses co-infect a cell, the replication enzymes they produce represent a common resource that must be produced by at least some of them (Turner & Chao, 2003). Similarly, adhesive polymers in bacteria (Rainey & Rainey, 2003) and invertase in yeast (Gore *et al.*, 2009) are public goods because they are diffused outside the cell; their production is costly, but their lack, if nobody produces them, can be lethal. The amoeba *Dictyostelium discoideum* when confronted with starvation initiates a developmental process that leads to the formation of a ball of spores, which reproduce, supported on a stalk, whose cells die. If one individual has the capacity to avoid the stalk cell fate, it may have a selective advantage; however, some individuals must volunteer to produce the stalk (Bonner, 2008). A similar

Correspondence: Marco Archetti, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138-2902, USA.  
Tel.: +1 617 496 8146 e-mail: archetti@fas.harvard.edu

situation occurs in *Myxococcus xanthus*, where cells become either spores or nonspores (Velicer *et al.*, 2000). In vertebrates, in groups that rely on alarm calls as a defence against predators, one individual is enough to give the alarm, and the alarm call may have nonnegligible costs (Searcy & Nowicki, 2005); predator attack, however, if nobody gives the alarm, can be lethal. In general, in all the cases mentioned above, each individual prefers to avoid the cost of volunteering and exploit the benefit of the public good, but someone must volunteer and pay the cost of producing the good; if nobody volunteers, the cost paid is greater than the cost of volunteering.

### Volunteer's dilemma

The volunteer's dilemma, as first proposed in the social sciences (Diekmann, 1985), is an N-person game in which a public good is produced if and only if at least one player volunteers to pay a cost. The basic model is the following: N persons are sentenced to jail for 10 years unless at least one of them volunteers to confess; in this case, the one that confesses (the volunteer) is imprisoned for 1 year, and the others are released. Clearly, volunteering produces a common good, but the action is costly to the volunteer himself; the volunteer benefits from his action if nobody else volunteers, but the cost of his action is wasted if someone else volunteers; this is the dilemma (note that, although the story involves prisoners, the game is rather different from the PD). In the absence of coordination should one volunteer? And how often? The production of a collective good by the contribution of a single volunteer is inefficient without a coordination mechanism for selecting who volunteers.

The volunteer's dilemma can be applied to many cases in biology in which one or few volunteers are enough to produce a common good, but volunteering is costly. I analyse a model of the volunteer's dilemma extending it to cases of relatedness between group members and to situations in which more than one volunteer is necessary to produce the public good. Because the PD predicts that defection is the stable equilibrium (in the absence of iterations), the typical question that arises in public goods games is: why does cooperation persist in nature? I show that in the volunteer's dilemma, cheaters and cooperators coexist in a mixed equilibrium. The real dilemma is not why cheaters do not invade, but how to increase the probability that the public good is produced. I suggest that a practical solution is brinkmanship, the deliberate increase of risk.

## The model

### The basic model

In the basic model, each of  $N$  individuals can choose to volunteer (*Volunteer*) or not (*Ignore*). A public good is

produced if and only if at least one individual volunteers. Volunteering has a cost  $c > 0$  (the individuals that volunteer have a payoff  $1 - c$  and the ones that do not have a payoff 1); if nobody volunteers, the public good is not produced and everybody pays a cost  $a > c$  (payoff  $1 - a$ ). Each individual, therefore, prefers that the public good is produced, but also prefers that it is someone else to volunteer.

If  $N = 2$ , the game with the two strategies *Volunteer* and *Ignore* has two asymmetric pure-strategy equilibria in which only one player volunteers, but they require coordination: it only works if the players decide in advance who is going to volunteer and when. The game has also a symmetric mixed-strategy equilibrium, which does not require coordination, in which  $1 - c = \gamma(1 - a) + (1 - \gamma)(1)$ , where  $\gamma$  is the probability of ignoring (not volunteering); therefore, at equilibrium  $\gamma_{\text{eq}} = c/a$ . This two-player game has the same structure and outcome as the hawk-dove (or chicken, or snow-drift) game (Maynard Smith & Price, 1973; Sugden, 1986; Doebeli & Hauert, 2005; Kun *et al.*, 2006). Importantly, however, the volunteer's dilemma is not an N-person version of the hawk-dove game with pairwise interactions. Interactions in the volunteer's dilemma are not pairwise but collective. The fitnesses of the pure strategies *Volunteer* ( $W_V$ ) and *Ignore* ( $W_I$ ) are:

$$W_V = 1 - c$$

$$W_I = \gamma^{N-1}(1 - a) + (1 - \gamma)^{N-1}$$

and the fitness of the mixed strategy is:

$$W_{\text{MIX}} = \gamma W_I + (1 - \gamma) W_V$$

The mixed-strategy equilibrium can be found by equating the fitness of the two pure strategies, which gives:

$$\gamma_{\text{eq}} = \left(\frac{c}{a}\right)^{1/(N-1)}.$$

This has interesting and counterintuitive consequences. First, and this is intuitive, the probability of ignoring increases with  $c$  and decreases with  $a$ . Second, the probability of ignoring increases with  $N$ : for example, with  $c = 0.3$  and  $a = 1$ ,  $\gamma_{\text{eq}} = 0.3$  with  $N = 2$ ;  $\gamma_{\text{eq}} = 0.87$  with  $N = 10$ ;  $\gamma_{\text{eq}} = 0.98$  with  $N = 50$ . The rationale is that when there are more players, each relies more on somebody else volunteering.

What is perhaps less intuitive is that the probability that *nobody* volunteers, and that therefore the public good is not produced, also *increases* with  $N$ . Because  $\gamma_{\text{eq}} = (c/a)^{1/(N-1)}$  then  $\gamma_{\text{eq}}^{N-1} = c/a$ , and  $\gamma_{\text{eq}}^N = \gamma_{\text{eq}} c/a$ , which is increasing in  $N$  (because  $\gamma_{\text{eq}}$  is increasing in  $N$ ). If  $a = 1$  and  $c = 0.3$ , for example, the probability that nobody volunteers is  $\gamma_{\text{eq}}^2 = 0.09$ ; with  $N = 10$ , the probability rises to  $\gamma_{\text{eq}}^{10} = 0.26$  and with  $N = 50$  to  $\gamma_{\text{eq}}^{50} = 0.29$ . Therefore, the more individuals are available to volun-

teer, the *less* likely it is that someone will actually volunteer. Increasing group size does not lead to a beneficial effect for the group.

**Relatedness**

In many situations of biological interest, individuals in a group are genetically related to some extent. How does relatedness affect the volunteer’s dilemma? If  $r$  is the average relatedness with other members of the group, the inclusive fitness of the two pure strategies *Volunteer* ( $W_V$ ) and *Ignore* ( $W_I$ ) can be written as:

however, simplifies calculations when more than one volunteer is needed – and will be more useful in the next section). The relative fitness ( $W$ ) of a player and that of the other group members ( $W_o$ ) written as a function of his own probability of volunteering ( $\gamma$ ) and the probability that other individuals in the group volunteer ( $\gamma_o$ ) are:

$$W = \gamma[\gamma_o^{N-1}(1 - a) + (1 - \gamma_o^{N-1})(1)] + (1 - \gamma)(1 - c)$$

$$W_o = \gamma_o[\gamma_o^{N-2}(1 - a) + (1 - \gamma_o^{N-2})(1)] + (1 - \gamma_o)(1 - c)$$

$$W_V = \underbrace{\gamma^{N-1}(1 - c)}_{\text{nobody volunteers}} + \underbrace{(1 - \gamma^{N-1})(1 - c)}_{\text{somebody volunteers}} + r \left\{ \underbrace{\gamma^{N-1}(N - 1)(1)}_{\text{nobody volunteers}} + \underbrace{\sum_{i=0}^{N-2} f_i \left[ \underbrace{(N - 1 - i)(1 - c)}_{\text{volunteer}} + \underbrace{i(1)}_{\text{ignore}} \right]}_{\text{somebody volunteers}} \right\}$$

$$W_I = \underbrace{\gamma^{N-1}(1 - a)}_{\text{nobody volunteers}} + \underbrace{(1 - \gamma^{N-1})(1)}_{\text{somebody volunteers}} + r \left\{ \underbrace{\gamma^{N-1}(N - 1)(1 - a)}_{\text{nobody volunteers}} + \underbrace{\sum_{i=0}^{N-2} f_i \left[ \underbrace{(N - 1 - i)(1 - c)}_{\text{volunteer}} + \underbrace{i(1)}_{\text{ignore}} \right]}_{\text{somebody volunteers}} \right\}$$

where

$$f_i = \binom{N - 1}{i} \gamma^i (1 - \gamma)^{N-1-i}$$

is the probability that  $i$  of the other  $N - 1$  individuals (other than self) do not volunteer. This can be explained as follows: *Volunteer* has a direct payoff  $(1 - c)$  irrespective of the probability that someone else volunteer, whereas *Ignore* has a direct payoff 1 if somebody else volunteer, which happens with probability  $(1 - \gamma)^{N-1}$ , and  $(1 - a)$  if nobody else volunteer, which happens with probability  $\gamma^{N-1}$ . In addition, if nobody of the other  $N - 1$  members of the group (whose average relatedness to the focal individual is  $r$ ) volunteers, which happens with probability  $\gamma^{N-1}$ , their payoff is 1 if the focal individual plays *Volunteer*, and  $(1 - a)$  if the focal individual plays *Ignore*. If  $i$  of these  $N - 1$  individuals play *Ignore*, instead, which happens with probability  $f_i$ , the payoff for those  $(N - 1 - i)$  that volunteer it is  $(1 - c)$  and the payoff for the  $i$  that ignore is 1, irrespective of the strategy of the focal individual. The mixed equilibrium can be found by equating  $W_V$  and  $W_I$ .

The same result can be found using an alternative inclusive fitness approach (the payoff equating approach,

and the mixed equilibrium can be found by calculating:

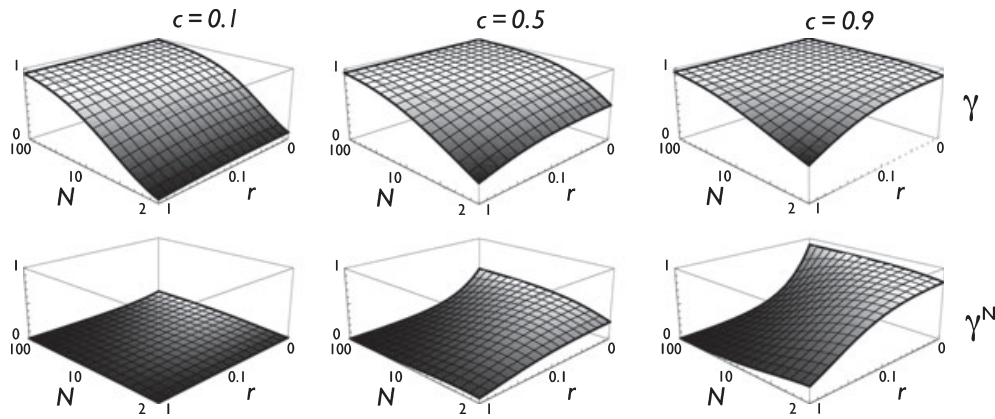
$$\frac{\partial W}{\partial \gamma} + (N - 1)r \frac{\partial W_o}{\partial \gamma} = 0.$$

Both approaches give:

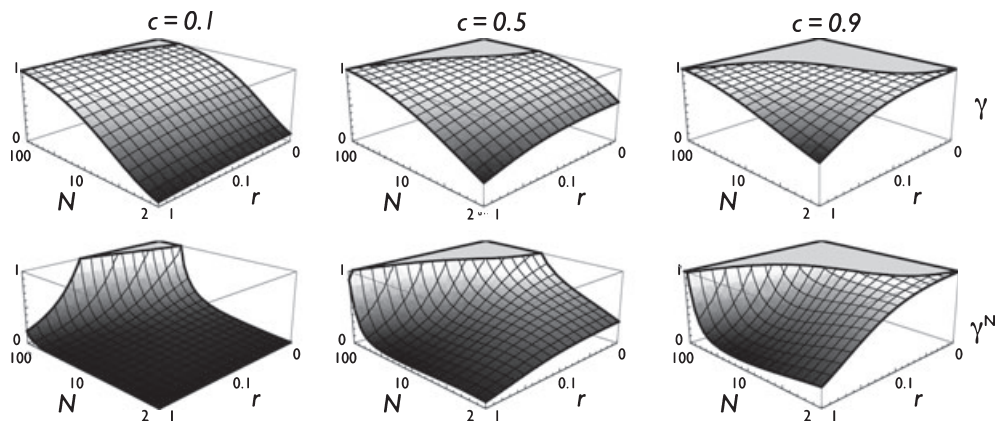
$$\gamma_{eq} = \left\{ \frac{c}{a[1 + r(N - 1)]} \right\}^{1/(N-1)}.$$

Again, the probability  $\gamma_{eq}$  of ignoring increases with  $N$ . If  $r$  is close to 1, increasing group size always increases the probability that the public good is produced. If  $r$  is small, however, the probability that nobody volunteers decreases with  $N$  only over a certain threshold, whereas an initial increase in group size still reduces the probability that the public good is produced (Fig. 1). Note that the fitness of the mixed strategy does not change.

If the cost  $a$  of failing to produce the public good is shared among the members of the group (for example,  $a = \alpha^N$ , with  $\alpha < 1$ ; Fig. 2), in most cases  $\gamma_{eq}^N$  (the probability that the public good is not produced) decreases with  $N$ , reaches a minimum and then increases again. Therefore, there is an optimal, intermediate value of  $N$  for which the probability that nobody volunteers is minimized.



**Fig. 1** The probability of not volunteering ( $\gamma$ ) and the probability that nobody volunteers ( $\gamma^N$ ) at equilibrium as a function of relatedness ( $r$ ) and of group size ( $N$ ).  $a = 1$ .



**Fig. 2** The probability of not volunteering ( $\gamma$ ) and the probability that nobody volunteers ( $\gamma^N$ ) at equilibrium as a function of relatedness ( $r$ ) and of group size ( $N$ ).  $a = 0.95^N$ .

**More than one volunteer**

In many cases, more than one volunteer will be necessary for the production of the public good. In principle, it would be realistic to model a public good whose magnitude increases smoothly with the number of volunteers, but in practice it makes little difference (results not shown here) to model the public good as a step function – that is, to assume that the public good is produced if and only if at least  $k$  individuals volunteer. In this case, the fitness of the two strategies is:

$$W_V = \sum_{i=0}^{N-1} f_i(1 - a_V) - c$$

$$W_I = \sum_{i=0}^{N-1} f_i(1 - a_I)$$

and

$$a_I = \begin{cases} 0 & \text{if } (N - 1) - i \geq k \\ a & \text{if } (N - 1) - i < k \end{cases}$$

$$a_V = \begin{cases} 0 & \text{if } (N - 1) - i \geq k - 1 \\ a & \text{if } (N - 1) - i < k - 1 \end{cases}$$

where  $i$  is the number of individuals that do not volunteer and  $(N - 1)$  is the number of other individuals (apart from the focal individual); therefore,  $(N - 1) - i$  is the number of volunteers apart from the focal individual (who can be a volunteer or not); the cost  $a$  is the same for focal volunteers and focal nonvolunteers, but whereas volunteers pay it only when  $(N - 1) - i < k - 1$  (because they do volunteer and only need  $k - 1$  other volunteers), nonvolunteers (*Ignore*) pay it when  $(N - 1) - i < k$ , because they need  $k$  others to volunteer.

The mixed-strategy equilibrium can be found by equating  $W_V$  and  $W_I$ . Because  $a_I = a_V$  for  $(N - 1) - i \geq k$  and for  $(N - 1) - i < k - 1$ , the only case in which  $a_I$  and

$a_V$  are different is  $(N - 1) - i = k - 1$ , that is,  $i = N - k$ ; therefore, equating  $W_V$  and  $W_I$  is equivalent to equating

$$-c + f_{N-k}(1 - 0) = f_{N-k}(1 - a)$$

that is,

$$\frac{c}{a} = \left( \frac{N - 1}{N - k} \right) \gamma^{N-k} (1 - \gamma)^{k-1}$$

With relatedness, the fitness functions become:

$$W_V = \underbrace{\sum_{i=0}^{N-1} f_i(1 - a_V)}_{\text{self}} - c + r \left\{ \underbrace{\sum_{i=0}^{N-1} f_i}_{\text{kin}} \left[ \underbrace{(N - 1 - i)(1 - a_V - c)}_{\text{volunteer}} + \underbrace{i(1 - a_V)}_{\text{ignore}} \right] \right\}$$

$$W_I = \underbrace{\sum_{i=0}^{N-1} f_i(1 - a_I)}_{\text{self}} + r \left\{ \underbrace{\sum_{i=0}^{N-1} f_i}_{\text{kin}} \left[ \underbrace{(N - 1 - i)(1 - a_I - c)}_{\text{volunteer}} + \underbrace{i(1 - a_I)}_{\text{ignore}} \right] \right\}$$

The mixed equilibrium can be found, as before, by equating the payoffs of the pure strategies. Again, because  $a_I = a_V$  for  $(N - 1) - i \geq k$  and for  $(N - 1) - i < k - 1$ , the only case in which  $a_I$  and  $a_V$  are different is  $i = N - k$ ; therefore, equating  $W_V$  and  $W_I$  is equivalent to equating

$$-c + f_{N-k}(1 - 0) + r f_{N-k}[(N - 1 - i)(1 - 0 - c) + i(1 - 0)]$$

and

$$f_{N-k}(1 - a) + r f_{N-k}[(N - 1 - i)(1 - a - c) + i(1 - a)]$$

that is:

$$\frac{c}{a[1 + r(N - 1)]} = \left( \frac{N - 1}{N - k} \right) \gamma^{N-k} (1 - \gamma)^{k-1}$$

The probability that the public good is not produced is:

$$\pi = \sum_{i=N-k+1}^N \binom{N}{i} \gamma_{\text{eq}}^i (1 - \gamma_{\text{eq}})^{N-i}$$

**Brinkmanship**

With  $k = 1$ , increasing  $a$  increases the probability that an individual volunteers ( $\gamma_{\text{eq}}$ ) and that the public good is produced; clearly  $\text{Lim}_{a \rightarrow \infty} \gamma_{\text{eq}} = 0$ . With  $k > 1$ , the effect of increasing  $a$  is the same, and it is stronger for intermediate values of  $k$  (Fig. 3). The highest probability that the public good is not produced ( $\pi$ ), and therefore the worst outcome for the group, occurs at intermediate values of  $k$ . It is important to notice that increasing  $a$  not only increases the probability that the public good is produced, but also fitness. The fitness ( $w_{\text{MIX}}$ ) of the mixed strategy at equilibrium clearly does not increase with  $a$  if  $k = 1$  (it remains equivalent to the fitness of *Volunteer*  $1 - c$ ); with  $k > 1$ , however, it does increase with  $a$  (the fitness of

*Volunteer*, and therefore the fitness of the mixed strategy, does change with  $\gamma$ ). In social dilemmas of this kind, therefore, brinkmanship (increasing the cost paid if the public good is not produced) is an optimal strategy both for the group and for the individual.

Moreover, fitness is reduced at intermediate values of  $k$ , at least for low values of relatedness, whereas with high relatedness it has a minimum for the highest values of  $k$  (Fig. 4). There are two effects to consider here. First,

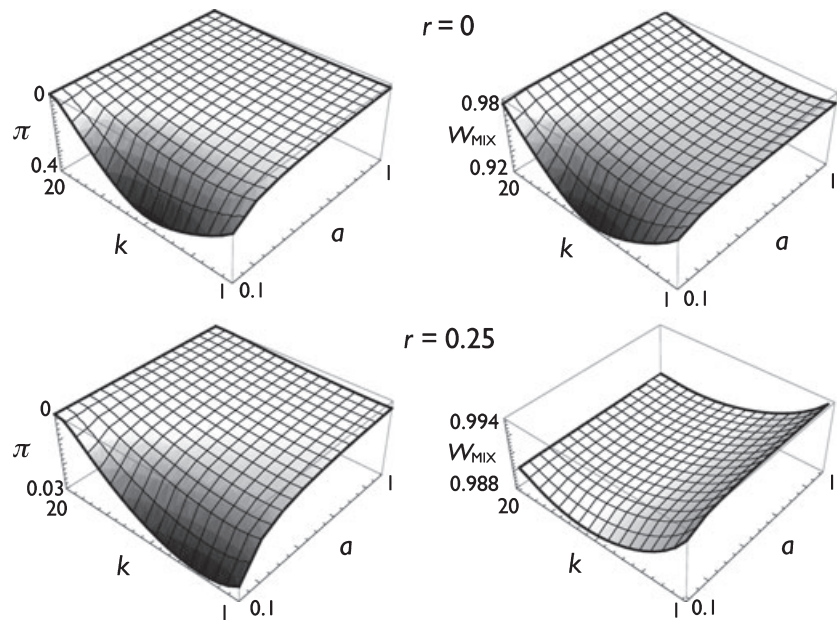
the probability that one's volunteering is wasted because not enough other individuals volunteer and the number of necessary volunteers is not reached; this increases with  $k$  (and is 0 with  $k = 1$  because if one volunteers in this case the public good is always produced). Second, the probability that one's volunteering is wasted because there are already enough volunteers; this decreases with  $k$  (and it is 0 if  $k = N$ , because then everybody knows that volunteering is necessary to produce the common good). These combined effects lead to the result that it is better, both for the individual and for the group, when the number of volunteers required to produce the common good is either very low or very high; for intermediate values of  $k$ , it is more likely that the common good is not produced and fitness is also reduced.

**Discussion**

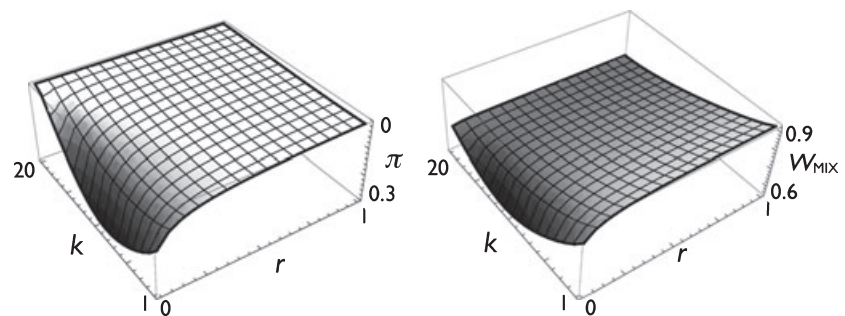
**Prisoners or volunteers?**

The tragedy of the commons (Hardin, 1968) is usually cited (Rankin *et al.*, 2007) as an analogy to explain why common resources are overexploited. Hardin (1968) used it as a metaphor to point out that individual self-interest does not necessarily lead to a benefit for the society, and that indeed in most situations Adam Smith's 'invisible hand' leads to a bad result for the society. His rebuttal was not a formal model but a verbal discussion based on the following example. Imagine a group of herders grazing cattle on common land; each herder only gains a benefit from his own flock, but when a herder adds more cattle to the land to graze, everyone shares the cost, which comes from reducing the amount of forage per cattle. Hardin (see also Rankin *et al.*, 2007) goes on stating that if the herders are driven only by economic





**Fig. 3** The probability that nobody volunteers ( $\pi$ ) and the fitness of the mixed strategy ( $W_{\text{MIX}}$ ) at equilibrium as a function of the number of volunteers required to produce the public good ( $k$ ) and of the cost paid when the public good is not produced ( $a$ ).  $N = 20$ ,  $c = 0.01$ ,  $r = 0$  (top) or  $r = 0.25$  (bottom). Fitness is normalized to  $W_{\text{MIX}}/[1 + (N - 1)r]$ .



**Fig. 4** The probability that nobody volunteers ( $\pi$ ) and the fitness of the mixed strategy ( $W_{\text{MIX}}$ ) at equilibrium as a function of relatedness ( $r$ ) and of the number of volunteers required to produce the public good ( $k$ ).  $N = 20$ ,  $a = 1$ ,  $c = 0.3$ .

self-interest, they will each realize that it is to their advantage to *always add another animal to the common*: they sacrifice the good of the group (by forgoing sustainable use of the resource) for their own selfish gain. Thus, herders will continue to add animals, eventually leading to a ‘tragedy’ in which the pasture is destroyed by overgrazing. This metaphor is useful to introduce the concept of social dilemmas, but can be misleading if the tragedy of the commons is equated to an N-person PD.

Note that, strictly speaking, the PD is a two-person game, and there is no such thing as an N-person PD without defining its structure, unless it means a game in which  $N$  individuals play the PD with pairwise interactions. Usually, however, an N-person PD is assumed to be the following game: individuals can be cooperators or defectors; cooperators pay a cost for contributing to the public good, whereas defectors refrain from doing so; after all individuals are given the chance to contribute to the public good, the accumulated contribution is multiplied by an enhancement factor, and the total amount

equally shared among all individuals (cooperators and defectors). Equating the tragedy of the commons to an N-person PD as defined above is misleading for the following reasons.

First, an N-person PD is a game, whereas the tragedy of the commons is a description of the equilibrium of a game: it simply means that the game is a social dilemma (the best strategy for the individual does not lead to the optimal outcome for the society). However, although it is true that an N-person PD leads to a Pareto inefficient equilibrium, there are other social dilemmas that are not an N-person PD. The volunteer’s dilemma is one example. The stag-hunt game (Skyrms, 2004) is another.

Second, most important, the fact that ‘it is to their advantage to always add another animal to the common’ (Hardin, 1968; Rankin *et al.*, 2007) is taken for granted, but it is by no means necessary. Imagine 10 herders with 10 cows each. If resources are depleted when 100 cows graze the pasture, each herder will introduce nine cows but will find it profitable to introduce the 10th only if at least one herder volunteers not to introduce his own

10th. It is not true that it is always an advantage to introduce one more cow, if we assume that 100 cows lead to complete resource depletion and that resource depletion is more costly than volunteering not to add the 10th cow. If, instead, the cost of complete resource depletion is irrelevant (smaller than the cost of not introducing the 10th cow), for example because it only affects future generations, introducing the 10th cow will always be profitable (as assumed by Hardin), but this would *not* be a social dilemma because resource depletion will be irrelevant for the current players. Either resource depletion is costly (more costly than volunteering not to introduce the 10th cow), and in this case we have a social dilemma like the volunteer's dilemma; or resource depletion is not costly, and in this case we have no social dilemma (as defined by game theory) at all. Note that, although resource depletion that affect future generations is indeed a tragedy for the society, it is not a social dilemma in the sense of game theory, because it does not affect the payoffs of the current players.

Rather than trying to understanding what kind of game Hardin had in mind, however, it is important to establish what kind of social dilemmas are relevant in biology. Although some situations in biology are likely to be N-person PDs, all the examples cited in the introduction are certainly more similar to a volunteer's dilemma than to an N-person PD. Some social dilemmas that are clearly volunteer's dilemmas have been classified as PD probably simply because it was assumed that any social dilemma is a PD. Some other cases have been classified as snowdrift game (SG), although in fact they are also volunteer's dilemmas because they do not involve pairwise interactions. Interactions between RNA phages co-infecting bacteria, for example, have been first described as PD (Turner & Chao, 1999) and subsequently as a SG (Turner & Chao, 2003). However, when viruses co-infect a cell, the replication enzymes they produce are a public good and interactions are collective, like in the volunteer's dilemma. Another clear and recent example of a case classified as SG and which is, instead, a volunteer's dilemma, is invertase production in yeast (Gore *et al.*, 2009). Collective hunting and territory defence in mammals, defined as a SG by Doebeli & Hauert (2005) are also volunteer's dilemmas. Sentinel behaviour, in not a SG (Doebeli & Hauert, 2005) but the asymmetric equilibrium of the volunteer's dilemma (different from the symmetric mixed equilibrium discussed here), which requires coordination.

### Biological volunteer's dilemmas

I have described a generalized model of the volunteer's dilemma, arguing that it applies to many cases of biological interactions. It is relevant for situations in which one or few individuals are enough to perform a costly action that produces a common good.

Microbes and social species often produce and consume resources that are costly, and each individual would find it more profitable to avoid the cost of producing them, but would find it more profitable to pay this small cost than to pay the larger cost that would occur if nobody produced the resource: replication enzymes for viruses co-infecting a cell (Turner & Chao, 2003), adhesive polymers in bacteria (Rainey & Rainey, 2003) and invertase in yeast (Gore *et al.*, 2009), as well as alarm calls in vertebrates (Searcy & Nowicki, 2005), are typical examples of diffusible, public goods. Collective breeding is another example, and even in the extreme case of *D. discoideum*, in which individuals that form the stalk die ( $c = 1$ ), the volunteer's dilemma can explain the existence of an intermediate number of volunteers; in this case a certain degree of relatedness between group members is necessary (which is the typical case in *Dictyostelium*). In general, however, volunteering does not require any relatedness.

It is important to point out that volunteering does not require any relatedness nor reciprocation. Each individual will volunteer with a certain probability for his own benefit. Relatedness, as we have seen, affects the results, but it is by no means essential. Reciprocation instead does not play any role here, although it would be interesting to model an iterated version of the volunteer's dilemma and see what happens in the repeated game.

In the volunteer's dilemma cheaters are maintained in a mixed equilibrium but they do not replace volunteers completely, because the complete lack of volunteers is more costly than the cost of volunteering. Therefore, each individual will volunteer with a certain probability; in alternative, a polymorphic population will exist with both cooperators and defectors. The problem of the evolution of cooperation, modelled as a volunteer's dilemma, is not to explain why cheaters do not invade; this is the usual question raised by public good games when they are modelled as an N-person PD, but it is not the case here. Cheaters, in the volunteer's dilemma, do invade and are maintained at the mixed equilibrium. The volunteer's dilemma, however, like the PD, leads to a disappointing result for the society, because the more individuals are available to volunteer, the less likely it is that someone actually volunteers and the public good is produced.

The fact that invasion by cheaters does not represent a problem for the volunteer's dilemma still leaves us with a problem: how to increase the probability that the public good is produced? This is a practical question for which a technical solution can be envisaged (Hardin, 1968), something that, when it involves human interactions, does not require changing our view of morality and social rules. In evolutionary games, it requires that a strategy that increases the benefit for the group has also an advantage for the individual, so that it can evolve by natural selection. My suggestion is that this can be achieved by brinkmanship, the deliberate increase of the

damage that occurs when the public good is not produced.

### Brinkmanship

Schelling (1960) has introduced the idea that players in situations of conflict may create a deliberate risk (brinkmanship) as a strategy to induce the other players to adopt a certain behaviour. Brinkmanship has been discussed mainly in the field of international relations, but it can be applied to cooperation and conflict among individuals.

Here, I apply it to the volunteer's dilemma. In this case, brinkmanship is achieved by increasing the damage that would result from not producing the public good. Increasing the cost  $a$  paid when the public good is not produced could be an effective evolutionary strategy to increase the level of cooperation because, as we have seen, it increases the frequency of volunteers, and because the cost  $a$  is shared by all members of the group. A mutant that induces a *higher* cost  $a$  will affect equally the fitness of all group members when the common good is not produced and therefore will not create differences in relative fitness among group members. On the other hand as we have seen, this will *increase* the probability that the common good is produced. Individuals in groups with higher  $a$  therefore will have higher fitness than individuals in groups with low  $a$ , and if they compete with each other, an increase of  $a$  will be favoured. Increasing the cost  $a$  paid when nobody volunteers, therefore, would be an effective strategy to increase cooperation among group members not only if enforced by an external authority, but also in evolutionary games, in which mutants for higher  $a$  could invade and go to fixation. Clearly, following an increase of this cost ( $a$ ), the probability of volunteering would not change immediately if it is genetically determined, but it will be adjusted over evolutionary times; if, instead, it is a rational response to a perceived risk (this might be the case in humans and other rational animals) it might change immediately or after a relatively rapid learning process.

Cases in which cooperation is favoured as a result of increasing the cost of cheating (or reducing the benefit of cheating) have been discussed especially in the social insects (Wenseleers & Ratnieks, 2006; Ratnieks & Wenseleers, 2008), but in these cases, the cost is an individual cost for cheating rather than a cost due to the failure to produce a public good. Brinkmanship as a strategy for public good games is probably more well-known in human social dilemmas. During the cold war, for example, being on the brink of disaster was the strategy to actually avoid a nuclear escalation (Schelling, 1960).

The concept of brinkmanship can have practical applications. Imagine, for example, a number of individuals discarding their waste in the environment, after which one or some of them can volunteer to pay a small

cost for cleaning up; if nobody does it, the resulting damage is greater than the cost of cleaning up. This is a volunteer's dilemma; a way to increase the probability that somebody does the cleaning up would be to *increase* the deleterious effects of the waste. This strategy might not sound appealing to a public authority, but it is perfectly rationale, and in effect does not require to be enforced by an external authority; it could be achieved by a single individual that made its waste more toxic (increased  $a$ ), provided this was common knowledge among players, and provided the right costs and benefits ( $a > c$ ) exist. It must also be the case that these individuals, after having the opportunity to cooperate with others at a local scale, compete with others on a more global scale.

This idea, that a higher risk induces higher cooperation is the concept of brinkmanship applied to public goods games. The rationale is that increasing the cost paid when the public good is not produced makes volunteering more likely. It is the same rationale behind the (counterintuitive) result that the probability that someone volunteers decreases with group size.

Reducing the cost  $c$  paid by volunteers, instead, does not seem an effective strategy. Obviously, it leads to an increase of volunteering. However, a mutant individual with a lower  $c$  would create an asymmetry in the group, and he would be the first to volunteer, because volunteering would be less costly for him. The presence of asymmetries in the costs or benefits is usually recognized as the solution to the volunteer's dilemma in the social sciences (Nalebuff & Bliss, 1984; Weesie, 1993). Mutants with a low  $c$ , however, would not invade because they would always be the ones that pay the cost of volunteering and have a lower fitness (unless having a lower  $c$  implies that these mutants do better in other situations, but then this case involves fitness effects beyond cooperation, which is not interesting for our discussion). Reducing  $c$ , therefore, although it could be enforced by an external authority, cannot be an evolutionary strategy to increase the probability that the public good is produced.

Another possible strategy for increasing the probability that the public good is produced is setting the number of individuals required to produce the public good ( $k$ ) to the optimal value. As we have seen, this value is usually the lowest or the highest, and never intermediate. This strategy, however, must be enforced by an external authority. One possible biological example is the following. Imagine a mycorrhizal association in which individual fungi must cooperate among themselves to maintain a symbiosis with a tree (fungi receive carbohydrates and in exchange the tree manages to absorb more water and nutrients from the soil; Kiers & Denison, 2008). If the tree reacts to the average amount of nutrients, cutting its supply of carbohydrates to the fungi unless a certain amount of nutrients is absorbed, the fungi play a volunteer's dilemma (among themselves; their coopera-



tive behaviour among themselves is not the same as the mutualism established with the tree). In this case, the tree can actually act as an external authority. By deliberately increasing his dependence on the fungi, that is the amount of nutrients required from the symbiosis, the tree can in effect increase  $k$  and by doing so increase the probability that the fungi cooperate and the common good is produced. Examples like this, however, are probably rare in nature and optimizing  $k$  is probably a more relevant strategy for human interactions.

Increasing the damage suffered when the public good is not produced (brinkmanship), instead, seems a practical, efficient strategy to increase cooperation. It increases both individual fitness and the benefit for the group, and does not require any enforcement by an external authority. This is a rather surprising solution for a social dilemma. Hardin (1968) suggested a technical solution for the problem of cooperation, one that should be enforced by an authority and which does not require changing our view of morality and altruism. The volunteer's dilemma suggests that, perhaps against our ideal of morality, one can remain selfish and actually increase the benefit for the society.

## Acknowledgment

This work was supported by NSF grant SES-0750480.

## References

- Axelrod, R. & Hamilton, W.D. 1981. The evolution of cooperation. *Science* **211**: 1390–1396.
- Bonner, J.T. 2008. *The Social Amoeba*. Princeton University Press, Princeton.
- Dawes, R.M. 1980. Social dilemmas. *Annu. Rev. Psychol.* **31**: 169–193.
- Diekmann, A. 1985. Volunteer's dilemma. *J. Conflict Resolut.* **29**: 605–610.
- Doebeli, M. & Hauert, C. 2005. Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecol. Lett.* **8**: 748–766.
- Gore, J., Youk, H. & van Oudenaarden, A. 2009. Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**: 253–256.
- Hardin, J. 1968. The tragedy of the commons. *Science* **162**: 1243–1248.
- Kiers, E.T. & Denison, R.F. 2008. Sanction, cooperation, and the stability of plant-rhizosphere mutualisms. *Annu. Rev. Ecol. Evol. Syst.* **39**: 215–236.
- Kun, A., Boza, G. & Scheuring, I. 2006. Asynchronous snowdrift game with synergistic effect as a model of cooperation. *Behav. Ecol.* **17**: 633–641.
- Luce, R.D. & Raiffa, H. 1957. *Games and Decisions: Introduction and Critical Survey*. Wiley & Sons, New York.
- Maynard Smith, J. & Price, G.R. 1973. The logic of animal conflict. *Nature* **246**: 15–18.
- Nalebuff, B. & Bliss, C. 1984. Dragon-slaying and ballroom dancing: the private supply of a public good. *J. Public Econ.* **25**: 1–12.
- Nowak, M. 2006. Five rules for the evolution of cooperation. *Science* **314**: 1560–1565.
- Rainey, P.B. & Rainey, K. 2003. Evolution of co-operation and conflict in experimental bacterial populations. *Nature* **425**: 72–74.
- Rankin, D., Bargum, K. & Kokko, H. 2007. The tragedy of the commons in evolutionary biology. *Trends Ecol. Evol.* **22**: 643–651.
- Ratnieks, F.L.W. & Wenseleers, T. 2008. Altruism in insect societies and beyond: voluntary or enforced? *Trends Ecol. Evol.* **23**: 45–52.
- Schelling, T.C. 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge.
- Searcy, W.A. & Nowicki, S. 2005. *The Evolution of Animal Communication*. Princeton University Press, Princeton.
- Skyrms, B. 2004. *The Stag Hunt and Evolution of Social Structure*. Cambridge University Press, Cambridge.
- Sugden, R. 1986. *The Economics of Rights, Cooperation and Welfare*. Blackwell, Oxford.
- Tucker, A. 1950. A two-person dilemma. In: *Readings in Games and Information* (E. Rasmusen, ed., 2001), pp. 7–8. Blackwell, Oxford.
- Turner, P.E. & Chao, L. 1999. Prisoner's dilemma in an RNA virus. *Nature* **398**: 441–443.
- Turner, P.E. & Chao, L. 2003. Escape from prisoner's dilemma in RNA phage  $\Phi 6$ . *Am. Nat.* **161**: 497–505.
- Velicer, G.J., Kroos, L. & Lenski, R.E. 2000. Developmental cheating in the social bacterium *Myxococcus xanthus*. *Nature* **404**: 598–601.
- Weesie, J. 1993. Asymmetry and timing in the volunteer's dilemma. *J. Conflict Resolut.* **37**: 569–590.
- Wenseleers, T. & Ratnieks, F.L.W. 2006. Enforced altruism in insect societies. *Nature* **444**: 50.

Received 7 June 2009; revised 5 August 2009; accepted 6 August 2009