PROCEEDINGS
OF
THE ROYAL
SOCIETY B

Proc. R. Soc. B (2009) **276**, 1167–1174
doi:10.1098/rspb.2008.1623

When does optional participation allow the evolution of cooperation?

Sarah Mathew* **and Robert Boyd***

Department of Anthropology, University of California, Los Angeles, CA 90095, USA

Altruistic punishment has been shown to invade when rare if individuals are allowed to opt out of cooperative ventures. Individuals that opt out do not contribute to the common enterprise or derive benefits from it. This result is potentially significant because it offers an explanation for the origin of large-scale cooperation in one-shot interactions among unrelated individuals. Here, we show that this result is not a general consequence of optional participation in cooperative activities, but depends on special assumptions about cooperative pay-offs. We extend the pay-off structure of optional participation models to consider the effects of economies and diseconomies of scale in public-goods production, rival and non-rival consumption of goods, and different orderings of the pay-offs of freeriding and opting out. This more general model highlights the kinds of pay-offs for which optional participation favours cooperation, and those in which it does not.

Keywords: evolution of cooperation; altruistic punishment; optional participation

1. INTRODUCTION

Humans cooperate much more than individuals of most other species. Among some other mammals, small groups of individuals cooperate to acquire mates, defend territories or raise offspring. Eusocial insects cooperate in large groups, but typically only with relatives. By contrast, large unrelated groups of humans cooperate and reap the vast range of benefits to be had from large-scale cooperation. Although the potential for cooperative social exchange is widespread, cooperative behaviour does not usually evolve because it is vulnerable to exploitation. Even if everyone benefits by behaving cooperatively, selection usually favours individuals who take the benefits without paying the cost, and, as a result, the immense benefit that can be generated for everyone through cooperation remains untapped.

While various mechanisms may have allowed the evolution of cooperation in other species (Hamilton 1963; Trivers 1971; Axelrod & Hamilton 1981), there is compelling evidence that the punishment of freeriders helps maintain human cooperation in large groups (Boyd *et al.* 2003). Experimental studies have shown that people are willing to suffer a cost to punish freeriders (Fehr & Gachter 2000; Henrich *et al.* 2006), and much research in anthropology has documented that the individuals who violate cultural norms are ostracized, gossiped about, forced to pay compensations or required to perform costly rituals of atonement. Such punishment is meted out even in societies without formal institutions of justice and policing.

There are two important questions regarding the evolution of punishment. First, how can punishment be evolutionarily stable? Punishment is costly to the individual who metes it out. While a population in which punishers are common can resist invasion by rare defectors, it can be invaded by second-order free riders who cooperate and thus avoid being punished, but who themselves do not incur the burden of sanctioning others. Plausible solutions to this problem have been proposed (Henrich & Boyd 2001; Milinski *et al.* 2002; Boyd *et al.* 2003; Panchanathan & Boyd 2004). The second question is, how can punishment increase when it is rare? When punishers are rare in the population, there are many defectors who must be punished, and so punishers suffer a large cost. As a result, in most models, selection does not favour punishment until the frequency of punishers exceeds a threshold value. Recently, however, Hauert *et al.* (2007) demonstrated that when participation in the cooperative endeavour is optional, not obligatory, punishers can invade even when they are rare, thus providing a potential explanation for the origins of altruistic punishment.

To understand the difference between optional and obligatory participation, consider the standard model of the evolution of cooperation in large groups. Groups of n individuals are randomly drawn from a large population. There are two strategies: *cooperators* contribute to a collective good and *defectors* do not. Each contribution benefits all members of the group whether or not they contribute. Participation is said to be obligatory because once the good is created, everyone in the social group inevitably benefits from it, whether they choose to or not. Many public goods fit this model well—no one can opt out of breathing cleaner air or the benefits of group defence. However, it is possible to opt out of other public goods—people *can* choose not to use a bridge even if it already has been built, or not to consume meat that has been brought into camp. Consequently, when participation is optional, in addition to the cooperator and defector strategies, there is a third strategy that opts out of playing the public goods game altogether. Such *loners* do not contribute to the production of the public good, nor do they consume its benefits. Instead, they obtain a fixed pay-off from a solitary pursuit.

Optional participation can radically alter the evolution of cooperation. Assume for a moment that there is no punishment. Then, if participation is obligatory, defectors

---

*Authors for correspondence (math.sarah@gmail.com; rboyd@anthro.ucla.edu).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2008.1623 or via http://journals.royalsociety.org.

Received 7 November 2008
Accepted 1 December 2008

always outcompete cooperators because they benefit from the collective good without incurring the cost. The unique, stable evolutionary outcome is a population with only defectors. By contrast, when individuals have the freedom to opt out, the population need not come to rest at an equilibrium of all defectors. Instead, the frequencies of the three strategies can cycle endlessly (Hauert *et al.* 2002). If in the absence of cooperators, the solitary activity pursued by loners yields a higher pay-off than defection, then a population of defectors will be invaded by loners. However, once loners are common, cooperators can thrive because there are no exploiters around. But once there are enough cooperators in the population, selection favours the exploiters.

Hauert *et al.* (2007) showed that when punishment of defectors is allowed and the population is finite, punishers rapidly invade this oscillating mix of cooperators, defectors and loners. Punishers are able to invade the oscillating mix because selection does not act against rare punishers during the loner or cooperator stages of the cycle, as defectors are also rare, and therefore punishment is cheap. Once punishers are frequent, they can stabilize cooperation by punishing rare defectors. Therefore, the choice to opt out provides an avenue for altruistic punishment to invade a population even when punishment is initially rare.

Here, we show that this conclusion depends on the particular assumptions shared by existing models of optional participation (Hauert *et al.* 2002, 2007; Fowler 2005; Brandt *et al.* 2006), and does not apply to many types of cooperation that characterize human societies. People cooperate in a wide range of activities. They work together to hunt the largest game in their habitats, mobilize large armies when being attacked, form raiding parties that plunder their weaker neighbours, work together to build shelters, roads, canals and defences, share food to reduce the risk from shortfall, participate in building government, enforce cultural norms, punish criminals, etc. While the pay-off structures of these activities differ in important ways, they can all be adequately approximated by the standard n-person public goods game when participation is obligatory. However, when participation is optional, the details of the pay-off structure matter. Using a more general model that can represent this range of activities, we show that only some forms of human cooperation can generate the cycling necessary for punishment to invade, while for many others there is no cycling and therefore punishment cannot increase when it is rare, even when participation is optional.

2. THE MODEL

In the standard, obligatory public-goods model, groups of n individuals are drawn from a very large population, and participate in a one-shot public goods game. There are two heritable strategies: cooperators contribute to the public good and defectors do not. Every individual in the group receives a benefit $b \cdot (\text{number of cooperators})/n$. Cooperators pay a cost c. n is the size of the group, which represents both the number of consumers and the number of contributors that maximize the social pay-off. An individual's reproductive success (either cultural or genetic) is proportional to his pay-off, and thus strategies with higher than average pay-off increase in frequency, and those with lower than average pay-off decrease.

This model can represent a wide range of real-world situations. Since the *per capita* pay-off increases with the number of cooperators, the standard model incorporates the idea that there are economies of scale in the production of public goods. The magnitude of this effect is given by the parameter b. The standard model can also represent both rival and non-rival public goods. One individual's consumption of a rival public good reduces the amount available to others, while each consumer of a non-rival good obtains the same benefit, regardless of how many consumers there are. Examples of rival goods include cooperative hunting of small game where storage is possible, capital facilities such as roads or bridges that are subject to congestion, and the booty produced by collective raiding. The public goods game widely used in experimental economics also assumes rival consumption. Non-rival public goods include military defence, investments in environmental quality and killing of very large game in the absence of storage facilities. Because the number of consumers is constant in the standard model, it applies equally to either type of good.

Existing models of voluntary participation cannot represent the same range of situations. To allow voluntary participation, Hauert *et al.* (2007) added a strategy that opts out. These loners neither contribute to the public good, nor consume it, but instead pursue a solitary activity with a fixed pay-off. This addition led to a narrowing of the applicability of the model for two reasons.

(i) Allowing voluntary participation means that the number of consumers of the public good varies even if the group size is constant, and as a result the same model cannot represent both rival and non-rival goods, nor does it necessarily allow for economies of scale. Hauert *et al.* (2007) chose to model a rival good, and as in the standard model, each additional cooperator increases public-goods production by a constant amount. But since the number of consumers can vary, it does not necessarily follow that the *per capita* benefit increases with the number of producers. Hauert *et al.* chose a pay-off structure where the *per capita* benefit produced depends only on the *ratio* of the number of producers and consumers; the absolute number of cooperators is irrelevant. So, for example, a pair of cooperators in a group of loners can generate the same *per capita* public benefit as a group of n cooperators. To prevent a single cooperator in a group of loners from gaining the full benefits of cooperation, Hauert *et al.* assumed a lone cooperator or defector surrounded by loners behaves just as a loner.

(ii) Because there are three strategies, a single cost parameter is no longer sufficient. Pay-offs depend on two factors: 'social pay-offs' that depend on the behaviour of others in their group and 'individual pay-offs' that depend on an individual's strategy, but are independent of the group composition. In the standard game, the cost of cooperation, c, is the difference between the individual pay-off of a cooperator and a defector. With voluntary participation, the difference in individual pay-offs between defectors and loners, and between cooperators and loners must also be accounted. Here, Hauert *et al.* assumed that loners have a higher individual pay-off than defectors who get a larger individual pay-off than cooperators. As we will see, this assumption fits nicely with some real-world situations, but not others.

Here, we generalize the Hauert *et al.* model, so that it can represent the same range of situations as the standard obligatory participation model. We consider both rival and non-rival goods, allow for economies of scale and different orderings of the individual pay-offs. First, we describe the pay-offs when the public good is non-rival, and then when it is rival.

(a) *Non-rival goods*

Consider the pay-off of a focal individual when there are i cooperators, j defectors, and $n-1-i-j$ loners among the other $n-1$ individuals in the group. Then, if $i+j>0$, the pay-offs to cooperators (V_C), defectors (V_D) and loners (V_L) are

$$V_C = h_c + \left(\frac{i+1}{n}\right)b, \tag{2.1}$$

$$V_D = h_d + \left(\frac{i}{n}\right)b \tag{2.2}$$

and

$$V_L = h_l. \tag{2.3}$$

A single cooperator or defector in a group of $n-1$ loners receives a pay-off of h_l. The first term in each expression gives the individual pay-off of that strategy. The difference between the individual pay-offs of two strategies is the trade-off between pursuing these strategies. For example, h_d-h_c is the cost of cooperating, c, in the standard model. Existing voluntary participation models (Hauert *et al.* 2002, 2007; Fowler 2005; Brandt *et al.* 2006) assume that $h_l>h_d>h_c$. As we will discuss below, different orderings of the individual pay-offs of the three strategies are consistent with different cooperation problems in nature. The second terms in the expressions for V_C and V_D give the amount of public goods produced as a function of the number of cooperators in the group. We assume that $b>0$, which means that increasing the number of cooperators increases public-goods production. Because the good is non-rival, the benefit to each individual in the group is the amount of public goods produced. Thus, the *per capita* benefit from the public good also increases with the number of cooperators.

(b) *Rival goods*

Again, consider the pay-off of a focal individual, and let i and j give the number of cooperators and defectors among the other $n-1$ individuals in the group. Then, if $i+j>0$, the pay-offs of each of the three types are

$$V_C = h_c + \frac{\left(\frac{i+1}{n}\right)b + \left(\frac{i+1}{n}\right)^2 s}{i+j+1}, \tag{2.4}$$

$$V_D = h_d + \frac{\left(\frac{i}{n}\right)b + \left(\frac{i}{n}\right)^2 s}{i+j+1} \tag{2.5}$$

and

$$V_L = h_l. \tag{2.6}$$

As before, a single cooperator or defector in a group of $n-1$ loners receives a pay-off of h_l. The first term in each expression gives the individual pay-off as in the non-rival case. The second terms in the expressions for V_C and V_D give the benefit to the individual from production of the public good. The numerators in these expressions give the amount of public goods produced, and the denominators the number of consumers of the good. We assume that $s+b>0$, which means that the public good produced by a group of n cooperators has a positive benefit. If $s=0$, the model reduces to the existing voluntary participation models; there are no economies of scale in production. If $s>0$, the *per capita* production of the public good increases as the number of cooperators increases; there are economies of scale in production. If $s<0$, increasing the number of cooperators decreases the *per capita* production. Because each additional contributor consumes the same amount, the *per capita* pay-off (i.e. production/number of consumers) increases with the number of contributors when there are economies of scale in production, remains constant when there are no economies or diseconomies of scale in production and decreases with the number of contributors when there are diseconomies of scale in production.

As in the non-rival case, the number of contributors is divided by n to capture the idea that the benefit produced per potential recipient by a cooperative act declines as n increases. Note that dividing by n does not mean the public good is divided by the number of consumers. The number of consumers is determined by the frequency of types in the population. The parameter n is determined by the underlying ecology that generates gains from cooperation in a particular context. To illustrate, consider a simplified version of a model of cooperative hunting, where the likelihood of hunting success increases linearly as more individuals contribute until there are n hunters. Now, if the cooperative venture in question requires 10 hunters to maximize the likelihood of capturing prey, then each hunter increases the likelihood of prey capture by 10 per cent. By contrast, if a large game hunt requires 100 hunters to maximize the success rate, then each hunter increases the likelihood of hunting success only by 1 per cent. This scaling is obtained by dividing the number of contributors by n.

To compute the average fitness of each strategy, we assume that the groups are drawn at random from a large population in which x, y and z are the frequencies of cooperators, defectors and loners. This means that i and j in the pay-off expressions are multinomial random variables with sample size $n-1$ and probabilities x and y, and standard methods can be used to compute the expectations (conditioned on the fact that $i+j>0$). The expressions for the expected pay-offs as a function of x, y and z are given in the online material.

An individual's fitness is the sum of the pay-offs they receive from the game and a baseline fitness value w_0. Their fitness determines the expected number of offspring they contribute to the next generation. Offspring always have the same behaviour type as the parent, and so, strategies whose behaviour yields a higher pay-off will increase in frequency in the population. Let p be the frequency of strategy A. Then the change in the frequency of A in one generation is

$$\Delta p = p\frac{W_A - \bar{W}}{\bar{W}}, \tag{2.7}$$

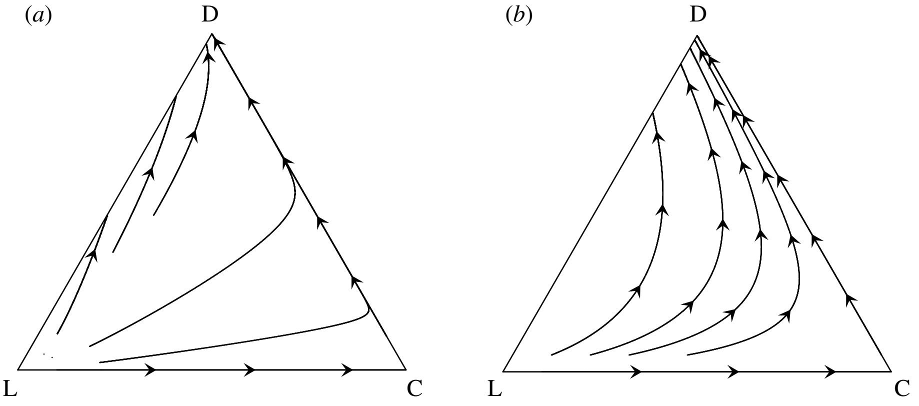

Figure 1. Oscillations do not occur when loners have the same pay-off as a defector in a group of all defectors. Instead, the loner–defector boundary is a stable attractor. (*a*) Non-rival public good. $b=10$, $n=10$, $h_c=1$, $h_l=2.5$, $h_d=2.5$, $w_o=200$. (*b*) Rival public good. $b=75$, $s=0$, $n=15$, $h_c=1$, $h_l=3$, $h_d=3$, $w_o=200$.

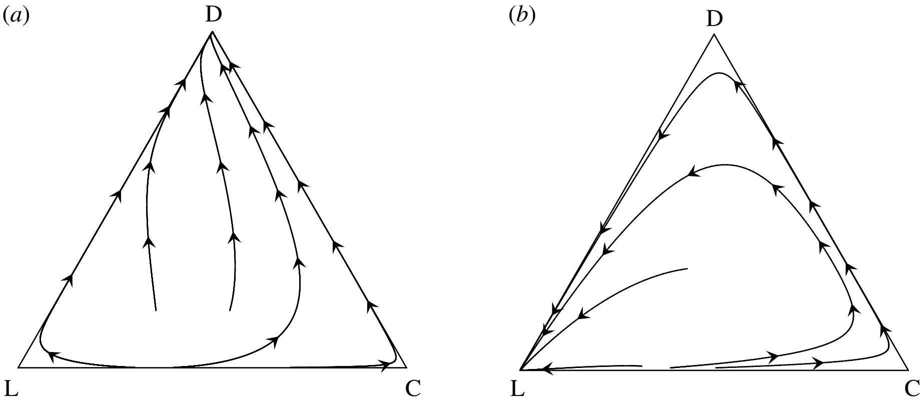

Figure 2. No oscillations occur when two cooperators in a group of loners have a lower *per capita* pay-off than a loner. Then there is an unstable equilibrium on the L–C boundary. The long-run equilibrium, for both rival and non-rival goods, is a population of all loners or all defectors depending on which has the higher individual pay-off. (*a*) Non-rival public good, $h_d > h_l$. $b=5$, $n=15$, $h_c=1$, $h_l=3$, $h_d=3.5$, $w_o=200$. (*b*) Rival public good with increasing returns to scale, $h_d < h_l$. $b=30$, $s=50$, $n=10$, $h_c=1$, $h_l=6$, $h_d=3$, $w_o=100$.

where W_A is the average fitness of individuals with strategy A and $\bar{W}$ is the average fitness of all strategy types in the population. We determined the stability of the pure equilibria analytically. To examine the evolutionary dynamics in mixed populations, we numerically iterated equation (2.7).

3. RESULTS

Stable oscillations occur only if the following three conditions are satisfied:

(i) loners have higher pay-off than a defector in a group of all defectors,
(ii) the pay-off of a cooperator in a group with another cooperator and the rest loners is greater than the pay-off of loners, and
(iii) the *per capita* pay-off from the public good does not decrease as the number of cooperators increases.

Here, we explain why oscillations in the frequencies of the three strategies occur when these conditions are satisfied.

(a) *Individual pay-off of loners must be greater than that of defectors*

In both the rival and non-rival cases, if $h_d \ge h_l$, there are no oscillations and a population of all defectors is the long-run outcome. When $h_d > h_l$, defectors invade any mix of loners and cooperators; the only stable equilibrium is a population of defectors. When $h_d = h_l$, the defector–loner boundary is an attractor, but points on the boundary are neutrally stable. In the long run, selection will take the population towards the pure defector equilibrium, because any disturbance away from the defector–loner boundary will favour defectors over loners. Figure 1 shows representative dynamics when $h_d = h_l$ for non-rival and rival goods. Thus, unless the *per capita* pay-off in a group of n defectors is less than the pay-off from the solitary pursuit, the optional participation model has the same outcome as the standard obligatory participation model, i.e. a population of all defectors.

(b) *Two cooperators in a group of loners must have a higher* per capita *pay-off than loners*

In the non-rival case, this requires that $h_c + (2b/n) > h_l$ and in the rival case that $h_c + (b/n) + (2s/n^2) > h_l$. When the frequency of cooperators is low, most of them are the only cooperator in their group, and thus they behave just as loners. Almost all of the rest of the cooperators are in groups with a single other cooperator. Thus, when the *per capita* pay-off that results from the public goods produced by two cooperators is less than the pay-off from the solitary pursuit, cooperators cannot invade a population of loners and therefore oscillations cannot

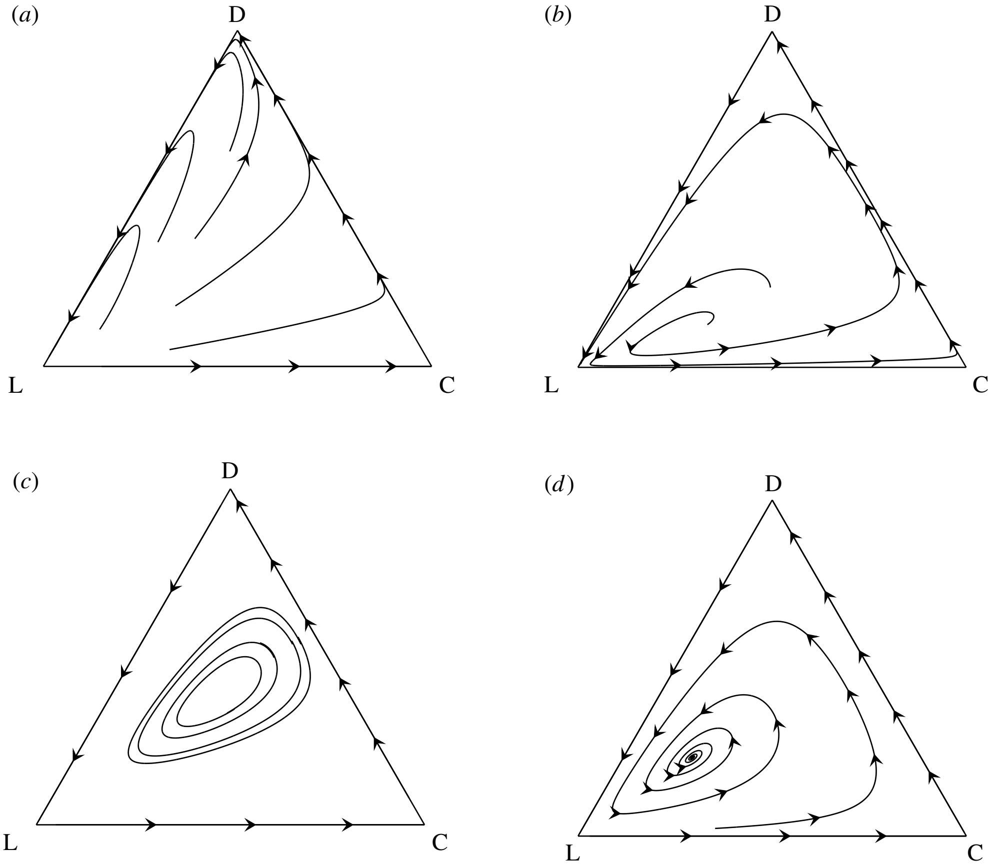

Figure 3. Oscillations occur when loners have a higher individual pay-off than defectors and two cooperators have a higher *per capita* pay-off than a loner. When there are increasing or constant returns, the oscillations persist. When there are decreasing returns they converge to a stable internal equilibrium. (*a*) Non-rival public good. $b=10$, $n=10$, $h_c=1$, $h_l=2.8$, $h_d=2.5$, $w_o=200$. (*b*) Rival good with increasing returns to scale. $b=50$, $s=50$, $n=10$, $h_c=1$, $h_l=6$, $h_d=3$, $w_o=100$. (*c*) Rival public good with constant returns to scale. $b=100$, $s=0$, $n=10$, $h_c=1$, $h_l=6$, $h_d=2.5$, $w_o=200$. (*d*) Rival public goods with decreasing returns to scale. $b=100$, $s=-40$, $n=10$, $h_c=1$, $h_l=6$, $h_d=3$, $w_o=100$.

occur (figure 2). The long-run outcome depends on the individual pay-offs: if $h_d \geq h_l$, then the long-run evolutionary outcome is a population of defectors. If $h_d < h_l$, then loners are the only evolutionary stable strategy.

(c) *The* per capita *pay-off from the public good must not decrease as the number of cooperators increases*

If loners have a higher individual return than defectors, and a pair of cooperators in a group of loners has a higher *per capita* benefit than a loner, then oscillations occur. The first condition allows rare loners to invade a population of defectors. The second condition allows cooperators to invade a population of loners. These conditions lead to long-run oscillations as long as the *per capita* pay-off from the public good does not decrease as the number of cooperators increases.

When consumption is non-rival (figure 3*a*), the *per capita* pay-off always increases as the number of cooperators increases. In this case, selection takes the population from any interior region where all three strategies are present towards the boundary. Once on the boundary where only two strategies are present, selection takes the population towards one of the pure equilibria. All of the pure equilibria are unstable. The pure defector equilibrium can be destabilized by a rare invading loner. The pure cooperator equilibrium can be invaded by a rare defector. The pure loner equilibrium can be destabilized by rare cooperators.

When the consumption is rival, the outcome depends on the returns to scale in production. When there are economies of scale ($s>0$), the oscillations spiral out to the boundary as in the non-rival case (figure 3*b*). When there are constant returns to scale ($s=0$), the oscillations occur as metastable closed orbits for a continuous-time model of replication. When selection is weak, the discrete time replication shown in figure 3*c* is a close approximation to a continuous-time model. In both of these cases, the population will continuously oscillate, and create the environment for altruistic punishment to invade as described in Hauert *et al.* (2007). When there are diseconomies of scale ($s<0$), the oscillations spiral inward to a stable polymorphic equilibrium at which all the three strategies are present (figure 3*d*). It seems likely that punishing strategies will not rapidly invade such a polymorphic equilibrium because defectors are present and thus punishers will have lower fitness than cooperators.

4. DISCUSSION

In the more general model of optional participation presented above, oscillations in the frequencies of cooperators, defectors and loners occur only when three conditions are satisfied.

(i) The individual pay-off of defectors is lower than that of loners. This means that all-round defection yields a lower pay-off than opting out, allowing loners to invade a population of defectors.

(ii) A cooperator paired with another cooperator in a group, where all of the remaining individuals are loners, has a higher pay-off than a loner. Otherwise, rare cooperators cannot invade a population of loners.

(iii) There are increasing or constant returns to scale. When increasing the number of cooperators lowers the *per capita* production of the public good, oscillations occur but converge to a stable, polymorphic equilibrium at which all three strategies are present.

Hauert *et al.* (2007) showed that altruistic punishment invades only when such oscillations occur. Thus, unless these conditions are satisfied, there is no cooperation at evolutionary equilibrium. In what follows, we describe what kinds of real-world situations are likely to satisfy these conditions, and what kinds are not. These examples are meant to be illustrative. In any real-world situation, costs and benefits must be measured empirically.

First, however, we want to point out that these results indicate that opting out of *consumption* of the public good is not the key feature of optional participation models. Note that the above three conditions apply to both rival and non-rival goods. In the non-rival case, the fact that loners do not consume the good does not affect the pay-offs of the other two types. So, opting out of consuming the good is irrelevant. What does matter is that two cooperators produce a *per capita* public benefit larger than the pay-off to be had from pursuing a solitary activity. Then rare cooperators can invade a population in which loners are common, and oscillations are possible. However, this raises the question: why cannot rare cooperators invade a population of defectors? After all, oscillations can occur only when the individual pay-off of defectors is lower than that of loners, and so cooperators should find it easier to invade a population of defectors than loners. True, for rival goods, the presence of defectors in their group lowers the pay-off of a pair of cooperators, but for non-rival goods, they do not. So, non-consumption is not the key. Rather, it is that in all optional participation models, a single cooperator in a group of all loners is assumed to act just as a loner and thereby get the same pay-off as loners, whereas a single cooperator in a group of all defectors still cooperates, and therefore does worse than defectors. In other words, what drives the oscillations is not *opting out* itself, but the assumption that cooperators can detect loners and thereby modify their own behaviour appropriately. If it is plausible to assume that loners can be detected, then we can potentially also modify the cooperator strategy so that it does not contribute unless there are x-1 others who do not opt out where x is the number of cooperators necessary for cooperation to yield higher pay-off than the solitary pursuit. This would allow cooperators to invade loners more easily. On the other hand, if it is implausible that loners can be detected, then rare cooperators cannot invade loner populations even if the above conditions are met.

(a) *When is all-round defection worse than opting out?*

A defector can have a lower pay-off than a loner when defectors suffer an opportunity cost that loners do not, and, as a result, defectors lose out when there are no cooperators to exploit. For example, suppose that hunting parties leave camp for the day in search of large game. Cooperators work hard and take risks; defectors hang back and as a result the more defectors in a group, the lower the success rate. If kills are consumed before returning to camp, defectors give up the benefits of solitary foraging in order to consume the benefits of the cooperative venture. Then, if the success rate of a group of all defectors is low enough, a defector in such a group will be worse off than those who stayed at home and tended their garden. Defectors may also do worse than loners if consumers of the public good are forced to pay a fee to join the cooperative venture. For instance, suppose that young men must go through a costly scarification ritual in order to join a raiding party. This would not motivate defectors to switch to cooperation: even if an individual scarred himself to be accepted into the raiding party, he would still benefit from standing in the back of the line when confronted with the enemy. But it does alter the relative pay-offs of loners and defectors. If warrior parties without any cooperators usually fail to acquire booty, defectors who go through costly scarification do worse than loners who avoid the ritual.

However, in many situations, the pay-off to all-round defection is the same as the solitary pay-off. Again consider cooperative hunting, but now suppose that the kill is brought back to camp, a common pattern in contemporary hunter–gatherer societies. The defectors and loners both forage alone, but defectors opportunistically scrounge from the kill that cooperators bring home. Here, a defector in a group of all defectors has the same individual pay-off as a loner because defectors do not need to forgo the solitary pursuit in order to have access to the meat. Similarly, suppose that individuals experience significant variance in daily foraging success, so that if n foragers agree to share what they acquire everyday, they reduce their daily variance in food intake, and thus increase their fitness. Now, if all n individuals defect on this agreement, and do not contribute to the shared pool, then they experience exactly the same daily variance in food intake as those who opt out of the food sharing contract in the first place.

In some situations, all-round defection may also be better than opting out. There are many examples of defection that reduce, but do not eliminate, the benefits from cooperation. For instance, defection could be attacking prey from too far away, not being the first in the battle line, or in watching but not playing with the children under your care. In these cases, defection reduces the pay-off from the cooperative venture, but may not eliminate it, and as a result, all-round defection could still yield a higher pay-off than opting out.

(b) *When do two cooperators fare better than a loner?*

There are some kinds of human cooperation in which the cooperation of two individuals can lead to a substantial increase in fitness. For instance, in activities such as group vigilance or cooperative food sharing for risk reduction, two individuals may fare better than the solitary individual, even if n individuals are needed to generate the maximum *per capita* returns. Two individuals, who take turns keeping vigil while they go out foraging, will benefit from turn taking. Three, may benefit even more. And so on, until some optimal group size. Similarly, two foragers who share food and thereby insure themselves against shortfall fare better than solitary individuals who do not share.

Regardless of what the optimal size of the sharing group is, if there is variance in foraging success, then some sharing will be better than no sharing.

However, many kinds of public goods may be profitable only when a certain participation level is reached. For instance, two hunters cooperatively pursuing large game may do worse than a solitary hunter who goes after smaller game, although a hunting party of 10 might come home with enough to feed the entire village. Or two warriors who set off to raid their neighbours may be decimated, even though a large raiding party could return with a lot of booty. Similar economies of scale characterize public goods such as the production and maintenance of shared investments, like roads, canals and bridges. Elaborate defensive structures commonly found in the archaeological record such as forts, moats, palisades and watch towers are also examples. In all these cases, cooperation is unlikely to be profitable unless there are enough individuals contributing to the common enterprise. Or, put another way, even in the absence of defectors, cooperators can have lower fitness than loners if there are too few individuals contributing to the production of the public good. Importantly, human cooperation is unique among mammals in the ability to solve such kinds of large-scale public goods problems.

(c) *Returns to scale in cooperative ventures*

In many kinds of cooperation, increasing the number of individuals who cooperate to produce the good also increases the *per capita* pay-off. For non-rival goods, as long as the total amount of the public good produced increases with the number of cooperators, the *per capita* pay-off will also increase. For instance, if the public good is detection of enemy intruders, then three sentinels guarding the village stockade are more likely to detect intruders than two, four better yet, and so on, until some optimal number of guards on duty. Similarly, the construction of capital facilities that are not subject to congestion such as fortifications are non-rival joint investments with positive gains from increased contribution. For rival goods too, *per capita* pay-off can increase as the number of cooperators increases, but the conditions are more stringent. Now, each additional cooperator must increase the amount of the public goods produced *per capita* because the amount produced is divided by the number of consumers. In other words, the production of the public good must exhibit economies of scale. Nonetheless, many kinds of rival public goods may satisfy this condition. For example, cooperative hunts that involve driving game off a cliff or into a trap, warriors raiding for booty, hunting of large game or the construction of capital facilities such as shelters, granaries, livestock enclosures or bridges and roads that may be subject to congestion will probably have increasing returns to scale, until some optimal number of contributors is reached.

Some types of cooperation may have decreasing returns to scale. For non-rival goods, decreasing the *per capita* production will occur only when increasing the number of cooperators above two decreases the total amount produced. For instance, if two hunters can trap a prey with ease, but a third hunter may alert the prey and facilitate its escape, then the addition of the third contributor lowers the total production, and so the *per capita* pay-off is lowered. For rival goods, diseconomies of scale in production will lead to decreasing returns in the *per capita* pay-off. For instance, three hunters may be more likely than two to successfully trap the prey, but the addition of the third hunter may not increase the likelihood of hunting success enough to offset the loss from dividing the prey in thirds rather than in half. In all these cases, the optimal group size is two—bigger groups lead to lower *per capita* pay-off even if there are no defectors who consume without producing. When there are constant returns, all group sizes return the same *per capita* pay-off, so at the individual level people should be indifferent about the size of the group. Examples of the constant returns case may be empirically hard to find in nature. It requires that the *per capita* pay-off increase when the second individual contributes, but any further contributors add to the public-goods production by exactly the same amount as they reduce it through consumption.

Note that it is unclear whether large-scale cooperation will evolve when the *per capita* returns are decreasing or constant, because larger groups of cooperators do not fare better than smaller groups. In the case of decreasing returns, individuals can increase their pay-off if they split up and work together in groups of two. If returns to scale are constant, as assumed in Hauert *et al.* (2007), there is also no advantage to cooperating in groups larger than two. Cooperation in both these contexts can be maintained within dyads through reciprocity instead of altruistic punishment. But constant or decreasing *per capita* pay-offs in large-scale cooperation may occur if individuals are constrained to interact in larger groups by some external factor. For example, if there is only one site at which a fishing weir can be built, or a well dug, then it will not be possible for the group to split into a number of smaller cooperative units. In such cases, a public-goods model with decreasing or constant returns may be applicable.

(d) *The many faces of cooperation*

In the opening line of *Anna Karenina*, Tolstoy famously wrote that 'Happy families are all alike; every unhappy family is unhappy in its own way'. In the world of obligatory participation, cooperation is like a happy family. The same model can be applied to a vast range of real-world social situations, and the answer is always the same. When the possibility of opting out is introduced, the situation is not so happy. Now, depending on the details of pay-offs, quite different outcomes are possible. To predict the long-run evolutionary outcome, you have to specify the costs of contributing, defecting and opting out, whether the good is rival or non-rival, and the details of the economies of scale of public-goods production. For the particular choices made by Hauert *et al.* (2007) the long-run evolutionary outcome is likely to be cooperation enforced by punishment. However, we have shown that other choices lead to non-cooperative outcomes, and argued that these versions of the model correspond to interesting real-world cases of cooperation. Since opting out is frequently possible, a single mechanism is unlikely to explain the origins of all types of large-scale cooperation, as the diversity of evolutionary outcomes in this model suggests. Instead, the empirical details of the types and variety of public goods pay-offs observed in nature will be crucial to understanding the evolution of the many faces of human cooperation.

We thank two anonymous reviewers and a board member for their useful suggestions. S.M. would like to thank Juergen Neubauer for his advice. We also thank Richard McElreath for providing the BARYPLOT software we used to make the figures S.M. was supported by an NSF Graduate Research Fellowship.

REFERENCES

Axelrod, R. & Hamilton, W. D. 1981 The evolution of cooperation. *Science* **211**, 1390–1396. (doi:10.1126/science.7466396)

Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. 2003 The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100**, 3531–3535. (doi:10.1073/pnas.0630443100)

Brandt, H., Hauert, C. & Sigmund, K. 2006 Punishing and abstaining for public goods. *Proc. Natl Acad. Sci. USA* **103**, 495–497. (doi:10.1073/pnas.0507229103)

Fehr, E. & Gachter, S. 2000 Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994.

Fowler, J. H. 2005 Altruistic punishment and the origin of cooperation. *Proc. Natl Acad. Sci. USA* **102**, 7047–7049. (doi:10.1073/pnas.0500938102)

Hamilton, W. D. 1963 The evolution of altruistic behavior. *Am. Nat.* **97**, 354–356. (doi:10.1086/497114)

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. (doi:10.1126/science.1070582)

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. (doi:10.1126/science.1141588)

Henrich, J. & Boyd, R. 2001 Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *J. Theor. Biol.* **208**, 79–89. (doi:10.1006/jtbi.2000.2202)

Henrich, J. *et al.* 2006 Costly punishment across human societies. *Science* **312**, 1767–1770. (doi:10.1126/science.1127333)

Milinski, M., Semmann, D. & Krambeck, H. J. 2002 Reputation helps solve the 'tragedy of the commons'. *Nature* **415**, 424–426. (doi:10.1038/415424a)

Panchanathan, K. & Boyd, R. 2004 Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432**, 499–502. (doi:10.1038/nature02978)

Trivers, R. L. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)