Probabilistic participation in public goods games

Tatsuya Sasaki[1,*], Isamu Okada[2] and Tatsuo Unemi[3]

[1]*Graduate School of Engineering,* [2]*Department of Business Administration, and* [3]*Department of Information Systems Science, Soka University, 1-236 Tangi-machi, Hachioji, Tokyo 192-8577, Japan*

Voluntary participation in public goods games (PGGs) has turned out to be a simple but effective mechanism for promoting cooperation under full anonymity. Voluntary participation allows individuals to adopt a risk-aversion strategy, termed loner. A loner refuses to participate in unpromising public enterprises and instead relies on a small but fixed pay-off. This system leads to a cyclic dominance of three pure strategies, cooperators, defectors and loners, but at the same time, there remain two considerable restrictions: the addition of loners cannot stabilize the dynamics and the time average pay-off for each strategy remains equal to the pay-off of loners. Here, we introduce probabilistic participation in PGGs from the standpoint of diversification of risk, namely simple mixed strategies with loners, and prove the existence of a dynamical regime in which the restrictions no longer hold. Considering two kinds of mixed strategies associated with participants (cooperators or defectors) and non-participants (loners), we can recover all basic evolutionary dynamics of the two strategies: dominance; coexistence; bistability; and neutrality, as special cases depending on pairs of probabilities. Of special interest is that the expected pay-off of each mixed strategy exceeds the pay-off of loners at some interior equilibrium in the coexistence region.

Keywords: evolutionary game theory; public goods games; voluntary participation; stable interior equilibrium

1. INTRODUCTION

In a free society, voluntary responsibility is, without doubt, the most important factor allowing people to enjoy public goods, such as social security, environmental resources and peace. Human beings are frequently responsible for unrelated strangers whom they may meet once or not at all, even when shouldering responsibility is costly for each individual. The emergence of such altruistic behaviour is puzzling from an evolutionary viewpoint, since natural selection works in favour of free-riders, i.e. those who take advantage of the public goods without contributing. This puzzle is well known as the *Tragedy of the commons* (Hardin 1968).

One of the theoretical and experimental models (Kagel & Roth 1995) for this situation is the public goods game (PGG) that represents a natural extension of the prisoner's dilemma to an arbitrary number of players (Boyd & Richerson 1988; Hauert & Szabó 2003). The game is characterized as follows (Dawes 1980): groups of cooperators outperform groups of defectors, but defectors always outperform cooperators in any mixed group. This situation can be formalized as follows. We consider a large population of players, where $N(\geq 2)$ players are occasionally chosen using random sampling. The strategy of each player, i.e. whether to contribute (cooperate) or not (defect), is specified in advance and does not depend on the composition of a sampled group. The pay-offs P_c for cooperators (C) and P_d for defectors (D) in an N-player

group with n_c cooperators are given by

$$P_c = rc\frac{n_c}{N} - c \quad \text{and} \tag{1.1}$$

$$P_d = rc\frac{n_c}{N}, \tag{1.2}$$

where r denotes the multiplication factor; c denotes the cooperative cost; and $1 < r < N$. For simplicity and without loss of generality, we set the cost $c = 1$ for the rest of this paper. The last inequalities $1 < r < N$ state the essence of the public goods situation. For these games, defection is the dominating strategy and classical and evolutionary game theories conclude that all players become defectors and their pay-offs reduce to zero.

Most recent studies on the problem of mutual defection have focused on punishment (Clutton-Brock & Parker 1995; Fehr & Gächter 2000, 2002; Sigmund *et al.* 2001; Gürerk *et al.* 2006) and reputation (Alexander 1987; Nowak & Sigmund 1998, 2005; Wedekind & Milinski 2000). These mechanisms to suppress defection require *a priori* individual identification, in order that defectors may be punished or lose their reputations. Our main interest is in a more fundamental model that does not require such pre-identification. In a previous study, Hauert *et al.* (2002*a*,*b*) have proposed voluntary PGGs as a simple extension to the PGG concept. In this extension, players can decide whether or not to participate in PGGs. Those unwilling to join PGGs are termed loners (L), and instead of participating, they prefer to rely on a small but fixed pay-off σ ($0 < \sigma < r - 1$) with the result that the members in a group of cooperators are better off than loners, but loners are better off than members in a group of defectors. Each player is sampled a number of times, and obtains an average pay-off that depends on the player's own predetermined strategy, as well as on the composition of

*Author for correspondence (tsasaki@soka.ac.jp).

the entire population, i.e. the relative frequencies of the three strategies (C, D and L). In the case that no co-player participates, it is assumed that the player has no other option than to play as a loner and thus obtains pay-off σ. Since $1 < r < N$, the three strategies form a rock–scissors–paper cycle: if most players cooperate, it would be profitable to defect; if defectors are prevalent, it would be better to abstain from the PGG; and if most players are loners, then small interaction groups form frequently, and the average pay-off for cooperators will be higher than that for defectors and loners. However, defectors always do better than cooperators in any mixed groups. This is an example of Simpson's paradox (Simpson 1951). The evolutionary process is of Red Queen type: an unending cycle of adjustments (Hauert *et al.* 2002*a*), for which the addition of loners cannot create an asymptotically stable interior equilibrium at which cooperators and defectors coexist (Hauert *et al.* 2002*a*,*b*, 2006*a*). Moreover, according to the results stated by these authors, the time averages of the pay-offs for the three different strategies are equal and reduce to the pay-off of loners σ and thus no one does better or worse than loners in the long term.

We now address the interesting issue of whether it is generally the case that systems with loners can neither stabilize the dynamics nor improve the average pay-off in the population. In this paper, we extend voluntary PGGs in a very intuitive way. We consider two mixed strategies: one is the combination of cooperators and loners and the other one is of defectors and loners. These strategies participate in the PGG with some probability and otherwise act as a loner.

In real life, whether to participate in unpromising enterprises may not be an all-or-nothing option but rather 'probabilistically' changing according to circumstances and risk orientation of the individual. Let us take investments in stocks as an example. In this case, cooperative participants are fair players who keep to regulations and create equal opportunities for profits as a common resource in a free market. Defectors, on the contrary, are unfair players who create opportunities for unreasonable gains by irregularities such as insider trading. However, in practice, few people violate regulations every time they make an investment. One of the important problems for both types of players is the asset allocation between high-risk high-return products such as stock investments and low-risk low-return investments like bank deposits. Those who prefer safety and rely entirely on the latter can be viewed as loners. Rational investors will analyse risk-return trade-offs and consider diversified investments according to their own tolerance for risk, employing the maxim 'Do not put all your eggs in one basket'. The current extension into mixed strategies with loners therefore seems worthwhile, in that it is applicable in practice.

In §2 we analyse the replicator dynamics of the two mixed strategies and show that this can recover all the basic evolutionary scenarios and, in particular, result in a coexistence regime in which the average population pay-off exceeds the loner's pay-off. We shall refer to this extended model as PGGs with 'probabilistic participation'.

2. PROBABILISTIC PARTICIPATION

The two types of mixed strategies considered are defined by

$$\tilde{C} = p_x C + (1 - p_x)L \quad \text{and} \tag{2.1}$$
$$\tilde{D} = p_y D + (1 - p_y)L, \tag{2.2}$$

where p_x and p_y are continuous-valued parameters in [0,1] and represent the rate of participation in the PGG. We call $\tilde{C}$ a 'potential' cooperator strategy and $\tilde{D}$ a 'potential' defector strategy. At first, each sampled player stochastically decides whether or not to enter the game with his or her participation rate. Subsequently, if participating, the player actualizes his or her own potential strategy (C or D) and plays the game, and if not participating, behaves like a loner. Now let $\tilde{x}$ be the relative frequency of $\tilde{C}$, $P_{\tilde{c}}$ and $P_{\tilde{d}}$ the expected pay-offs of $\tilde{C}$ and $\tilde{D}$, respectively, and $\bar{P}$ the average pay-off in the population: $\bar{P} = \tilde{x}P_{\tilde{c}} + (1 - \tilde{x})P_{\tilde{d}}$, where $1 - \tilde{x}$ means the relative frequency of $\tilde{D}$. Following evolutionary game theory which assumes that a strategy's pay-off determines the growth rate of its frequency within the population (Weibull 1995; Hofbauer & Sigmund 1998), we obtain the following replicator equations representing the evolution of $\tilde{x}$:

$$\dot{\tilde{x}} = \tilde{x}(P_{\tilde{c}} - \bar{P}) = -\tilde{x}(1 - \tilde{x})(P_{\tilde{d}} - P_{\tilde{c}}). \tag{2.3}$$

Both homogeneous states ($\tilde{x} = 0, 1$) are trivial equilibria of the dynamics. In PGGs with probabilistic participation, a variable z corresponding to the fraction of loners is given by $z(\tilde{x}) = (1 - p_x)\tilde{x} + (1 - p_y)(1 - \tilde{x})$. The variable z represents the expected frequency of loner-like behaviour.

Let us denote the average pay-offs of a cooperator and a defector as P_c and P_d, respectively. According to Hauert *et al.* (2002*b*), the advantage of a defector over a cooperator is

$$P_d - P_c = 1 + (r-1)z^{N-1} - \frac{r(1-z^N)}{N(1-z)} =: F(z). \tag{2.4}$$

This advantage depends only on z. In the open interval $(0, 1)$, $F(z)$ has no roots for $r \le 2$ and has exactly one root $\hat{z}$ for $r > 2$. In the case of the mixed strategies, the expected pay-offs $P_{\tilde{c}}$ and $P_{\tilde{d}}$ are given by $P_{\tilde{c}} = p_x P_c + (1 - p_x)\sigma$ and $P_{\tilde{d}} = p_y P_d + (1 - p_y)\sigma$, respectively. We thus obtain the advantage of a potential defector over a potential cooperator,

$$P_{\tilde{d}} - P_{\tilde{c}} = (\sigma(p_x - p_y) - (r-1)p_x)(1 - z^{N-1})$$
$$+ r\frac{p_x p_y}{1-z}\left(1 - \frac{1-z^N}{N(1-z)}\right) =: \tilde{F}(z). \tag{2.5}$$

The details of the calculation are shown in the electronic supplementary material. $\tilde{F}(z) = 0$ is the equilibrium condition. In particular, for $p_x = p_y$ (z is constant), equation (2.5) is reduced to $\tilde{F}(z) = (1-z)F(z)$, thus equation (2.3) is also $\dot{\tilde{x}} = -\tilde{x}(1-\tilde{x})(1-z)F(z)$. This means that when $p_x = p_y$, the evolutionary dynamics of relative frequencies of potential cooperators and defectors is equivalent to that of pure cooperators and defectors in voluntary PGGs restricted in constant z. The straightforward calculation gives $d\tilde{F}/d\tilde{x} > 0$ for $p_x < p_y$ and $d\tilde{F}/d\tilde{x} < 0$ for $p_x > p_y$ (see electronic supplementary material). The monotonicity of $\tilde{F}(z)$ makes it easy to classify the phase diagram on the parameter space (p_x, p_y).

(a) *Classification of evolutionary dynamics*
Let us define four regions (i)–(iv) covering (p_x, p_y)-space, by the pair of signs of $\tilde{F}(z(\tilde{x}))$ at each point of $\{\tilde{x} = 0, \tilde{x} = 1\}$: (i) $(\ge 0, \ge 0)$, (ii) $(< 0, > 0)$, (iii) $(\le 0, \le 0)$, and (iv) $(> 0, < 0)$.
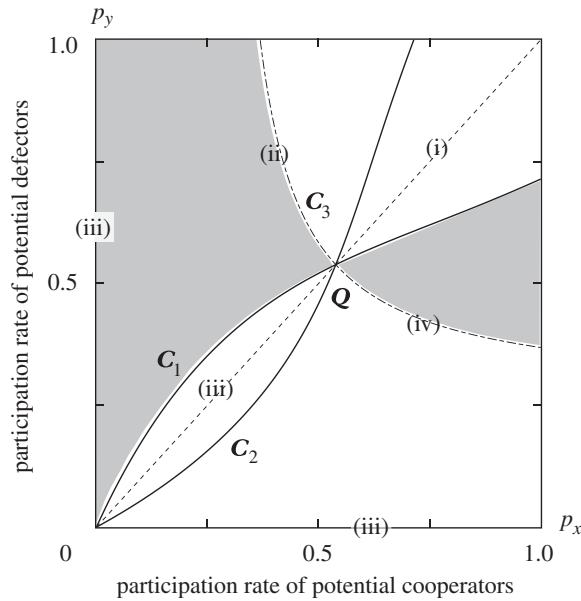
Figure 1. Phase diagram illustrating the different dynamical regimes ($N=5$, $r=3$, $\sigma=1$). The solid curves C_1 and C_2 separate (p_x, p_y)-space into the four regions (i)–(iv) characterized by the different dynamical regimes: (i) dominating potential defectors, (ii) coexistence, (iii) dominating potential cooperators, and (iv) bistability, except for the two-point set (v):$\{(0, 0), Q(1-\hat{z}, 1-\hat{z})\} = C_1 \cap C_2$, where the dynamics is neutral and $\hat{z}$ is the unique interior root of $F(z)$. The region (iii) includes the p_x- and p_y-axes. The interior equilibrium states satisfy $\bar{P} > \sigma$ in the grey-coloured part of (ii) and (iv) bounded by C_3 (the dashed curve), $\bar{P} < \sigma$ in the opposite part and $\bar{P} = \sigma$ on C_3.

The following features hold (figures 1 and 2; electronic supplementary material). The region (ii) satisfies $p_x < p_y$ and (iv) satisfies $p_x > p_y$. The diagonal $p_x = p_y$ is covered by (i) and (iii). For $r > 2$, the two boundaries, $C_1 : \tilde{F}(z(1)) = 0$ and $C_2 : \tilde{F}(z(0)) = 0$ are transversely crossing at $Q(1-\hat{z}, 1-\hat{z})$, where $\hat{z}$ is the unique interior root of $F(z)$. The four regions thus exist in any small neighbourhood of Q. We specifically denote the set of $(0, 0)$ and Q as (v).

Since $\tilde{F}(z)$ is monotone, for the regions (ii) and (iv), $\tilde{F}(z(\tilde{x}))$ has a unique interior root $\tilde{x}^*$, and for the regions (i) and (iii), no interior root. The five regions (i)–(v) then correspond to the five basic scenarios of evolutionary dynamics (Nowak & Sigmund 2004), which are as follows.

(i) $\tilde{D}$ *dominates* $\tilde{C}$ because $P_{\tilde{d}} > P_{\tilde{c}}$ holds. $\tilde{D}$-homogeneous state is globally stable.

(ii) If $\tilde{x} < \tilde{x}^*$, $\tilde{x}$ increases because $P_{\tilde{d}} < P_{\tilde{c}}$ holds. Conversely, if $\tilde{x} > \tilde{x}^*$, $\tilde{x}$ decreases because $P_{\tilde{d}} > P_{\tilde{c}}$ holds. As a result, $\tilde{C}$ and $\tilde{D}$ can invade each other and any mixed state converges to a unique interior equilibrium $\tilde{x}^*$. $\tilde{C}$ and $\tilde{D}$ *coexist* at globally stable $\tilde{x}^*$. Both trivial equilibria, $\tilde{C}$-homogeneous state and $\tilde{D}$-homogeneous state, are unstable.

(iii) $\tilde{C}$ *dominates* $\tilde{D}$ because $P_{\tilde{d}} < P_{\tilde{c}}$ holds. $\tilde{C}$-homogeneous state is globally stable.

(iv) If $\tilde{x} < \tilde{x}^*$, $\tilde{x}$ decreases because $P_{\tilde{d}} > P_{\tilde{c}}$ holds and then converges to $\tilde{x} = 0$. Conversely, if $\tilde{x} > \tilde{x}^*$, $\tilde{x}$ increases because $P_{\tilde{d}} < P_{\tilde{c}}$ holds and then converges to $\tilde{x} = 1$. Both $\tilde{C}$- and $\tilde{D}$-homogeneous states are asymptotically stable. The basins of attraction of these equilibria are divided by unstable $\tilde{x}^*$. In this *bistable*
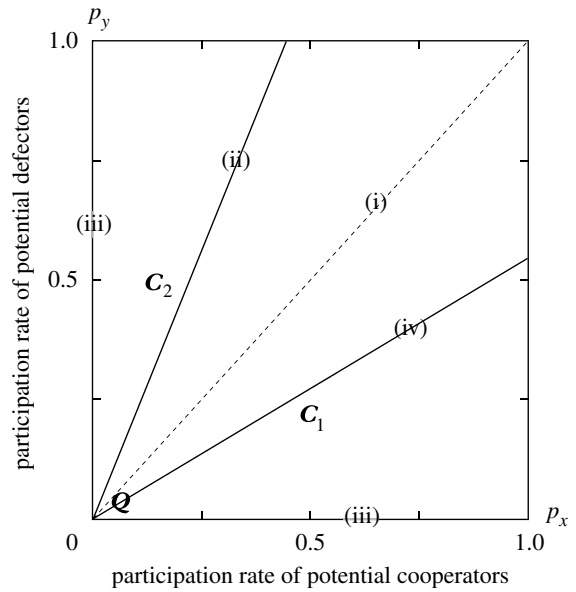
Figure 2. Phase diagram illustrating the different dynamical regimes ($N=2$, $r=1.5$, $\sigma=0.2$). The solid lines C_1 and C_2 separate (p_x, p_y)-space into the four different dynamical regimes (i)–(iv). The region (iii) consists of the p_x- and p_y-axes. $C_1 \cap C_2 = \{(0, 0)\}$. The point Q coincides with the origin. The interior equilibrium states sastisfy $\bar{P} < \sigma$ in (ii) and $\bar{P} > \sigma$ in (iv).

situation, the evolutionary fate of the system depends on the initial configuration.

(v) The fractions of $\tilde{C}$ and $\tilde{D}$ are unchanged because $P_{\tilde{c}} \equiv P_{\tilde{d}}$. The evolutionary dynamics is *neutral*.

A similar classification has been achieved by using the concepts of synergy and discounting for describing how benefits accumulate when multiple cooperators exist in an interaction group (Hauert *et al.* 2006b). Interestingly, these two models share almost the same arrangement of different dynamical regimes as is illustrated by the phase diagram (fig. 2 in Hauert *et al.* 2006b).

(b) *Average pay-off at equilibrium*
Given an interior equilibrium state, we evaluate the average population pay-off $\bar{P}^*$ in each moment, which is thus equal to the time average. $\bar{P}^* = \sigma$ defines a curve C_3 in (p_x, p_y). C_3 is included in the regions (ii) and (iv) and go through Q. C_3 divides each of (ii) and (iv) into two subregions, which are characterized by $\bar{P}^* > \sigma$ and $\bar{P}^* < \sigma$ (figure 1; electronic supplementary material). Therefore, for $r > 2$, these four types of subregions exist in any small neighbourhood of Q. But, for $r \leq 2$, the subregions $\bar{P}^* > \sigma$ in (ii) and $\bar{P}^* < \sigma$ in (iv) do not appear. We note that for $N=2$ (figure 2; electronic supplementary material), there is no chance of $r > 2$ owing to the precondition of the PGG $1 < r < N$. In the case of both $\tilde{C}$- and $\tilde{D}$-homogeneous states (except for the case that all members are actually pure loners), $\bar{P}^* > \sigma$ and $\bar{P}^* < \sigma$ hold, respectively.

3. SUMMARY AND DISCUSSION
Voluntary participation in PGGs (Hauert *et al.* 2002a,b), introducing the non-participating loners, results in oscillatory persistence of cooperation under full anonymity without requiring any cognitive abilities. However, this model also shows that the dynamics is unstable and in the

long run the average effect on the individual's pay-off is just the same as if all of the population were loners.

In this paper, we have discovered that these restrictions no longer hold for the probabilistic-participation strategies under suitable conditions. We consider two types of mixed strategies: potential cooperators and potential defectors who join the PGG with their own probabilities (participation rates) and otherwise act as loners. These strategies lead to very rich evolutionary dynamics including coexistence, depending on fixed pairs of participation rates. In the coexistence regime, stabilized cooperation (and defection) exactly arises in the form of a mixed strategy with loners. Moreover, when $N \geq 3$, $r > 2$ and some appropriate participation rates are given, this stable interior equilibrium enables the mixed strategies receive an average pay-off greater than the loner's pay-off every moment, which thus remains unchanged in the long run.

Interestingly, when $N = 2$, no coexistence state results in an average population pay-off greater than the loner's pay-off. For pairwise interactions, any coexistence state is thus taken over a homogeneous state of pure loners, and social interactions vanish, even supposing probabilistic participation. This outcome is equivalent to the results of previous research on the evolutionary Prisoner's Dilemma with voluntary participation (Hauert *et al.* 2002b; Hauert & Szabó 2005). However, if interacting groups are open to three or more players, a robust coexistence state against pure loners is possible. The openness to a third player is therefore a fundamental condition that makes freedom of participation to stabilize responsibility and to provide a stable pay-off, which is at least greater than the pay-off of a loner. Although it is uncertain whether the third player will sit down at the gaming table and what kind of participant (cooperator or defector) he or she might be, the vacant place has to be set to allow for the possibility of continuing social interactions.

Furthermore, it is also interesting to adopt other selection dynamics instead of considering the replicator dynamics. For example, the best-reply dynamics (Hauert *et al.* 2002a) would lead to convergence to oscillations between two points different in the dominating strategies, for the coexistence regime in the case of the replicator dynamics.

The most fascinating extension must be to introduce the process of mutation for the participation rate. The two mixed strategies occurring in this model contain the three pure strategies (C, D and L) corresponding to extreme values of the participation rate. Therefore, the addition of a mutation mechanism would bring about an intriguing question as to which kind of states would be adaptive, the rock–scissors–paper cycle or the stable coexistence state. In coevolution of more than one species, it has been much interesting to characterize the properties of fixed points at which there is no further phenotypic evolution (Marrow *et al.* 1996). Previous studies on two-species system have mainly dealt with examples of a discrete set of fixed points, but our model has a fixed 'curve', C_3 (figure 1), where mutational pressure on both p_x and p_y vanishes. Characterizing the properties of this curve would substantially contribute to solving of the above question.

REFERENCES

Alexander, R. D. 1987 *The biology of moral systems*. New York, NY: Aldine de Gruyter.

Boyd, R. & Richerson, P. J. 1988 The evolution of reciprocity in sizable groups. *J. Theor. Biol.* **132**, 337–356. (doi:10.1016/S0022-5193(88)80219-4)

Clutton-Brock, T. H. & Parker, G. A. 1995 Punishment in animal societies. *Nature* **373**, 209–216. (doi:10.1038/373209a0)

Dawes, R. M. 1980 Social dilemmas. *Annu. Rev. Psychol.* **31**, 169–193. (doi:10.1146/annurev.ps.31.020180.001125)

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

Fehr, E. & Gächter, S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)

Gürerk, Ö., Irlenbusch, B. & Rockenbach, B. 2006 The cometitive advantage of sanctioning institutions. *Science* **312**, 108–111. (doi:10.1126/science.1123633)

Hardin, G. 1968 The tragedy of the commons. *Science* **162**, 1243–1248. (doi:10.1126/science.162.3859.1243)

Hauert, C. & Szabó, G. 2003 Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interacton. *Complexity* **8**, 31–38. (doi:10.1002/cplx.10092)

Hauert, C. & Szabó, G. 2005 Game theory and physics. *Am. J. Phys.* **73**, 405–414. (doi:10.1119/1.1848514)

Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. 2002a Volunteering as red queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132. (doi:10.1126/science.1070582)

Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. 2002b Replicator dynamics for optional public goods games. *J. Theor. Biol.* **218**, 187–194. (doi:10.1006/jtbi.2002.3067)

Hauert, C., Holmes, M. & Doebeli, M. 2006a Evolutionary games and population dynamics: maintenance of cooperation in public goods games. *Proc. R. Soc. B* **273**, 2565–2570. (doi:10.1098/rspb.2006.3600)

Hauert, C., Michor, F., Nowak, M. A. & Doebeli, M. 2006b Synergy and discounting of cooperation in social dilemmas. *J. Theor. Biol.* **239**, 195–202. (doi:10.1016/j.jtbi.2005.08.040)

Hofbauer, J. & Sigmund, K. 1998 *Evolutionary games and population dynamics*. Cambridge, MA: Cambridge University press.

Kagel, J. H. & Roth, A. E. (eds) 1995 *The handbook of experimental economics*. Princeton, NJ: Princeton University press.

Marrow, P., Dieckmann, U. & Law, R. 1996 Evolutionary dynamics of predator prey systems: an ecological perspective. *J. Math. Biol.* **34**, 556–578. (doi:10.1007/s002850050021)

Nowak, M. A. & Sigmund, K. 1998 Evolution of indirect reciprocity by image scoring. *Nature* **339**, 573–577. (doi:10.1038/31225)

Nowak, M. A. & Sigmund, K. 2004 Evolutionary dynamics of biological games. *Science* **303**, 793–799. (doi:10.1126/science.1093411)

Nowak, M. A. & Sigmund, K. 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298. (doi:10.1038/nature04131)

Sigmund, K., Hauert, C. & Nowak, M. A. 2001 Reward and punishment. *Proc. Natl Acad. Sci. USA* **98**, 10 757–10 762. (doi:10.1073/pnas.161155698)

Simpson, E. H. 1951 The interpretation of interaction in contingency tables. *J. R. Stat. Soc. B* **13**, 238–241.

Wedekind, C. & Milinski, M. 2000 Cooperation through image scoring in humans. *Science* **288**, 850–852. (doi:10.1126/science.288.5467.850)

Weibull, J. W. 1995 *Evolutionary game theory*. Cambridge, MA: MIT press.