



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Stochastic evolutionary dynamics resolve the Traveler's Dilemma

Michael L. Manapat^{a,b}, David G. Rand^{a,c,d}, Christina Pawlowitsch^e, Martin A. Nowak^{a,f,g,*}^a Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA^b School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA^c Department of Psychology, Harvard University, Cambridge, MA 02138, USA^d Berkman Center for Internet and Society, Harvard Law School, Cambridge, MA 02138, USA^e Ecole d'économie de Paris, 75014 Paris, France^f Department of Mathematics, Harvard University, Cambridge, MA 02138, USA^g Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:

Received 7 September 2011

Received in revised form

27 February 2012

Accepted 14 March 2012

Available online 23 March 2012

Keywords:

Stochastic game theory
Evolutionary game theory
Evolutionary dynamics
Evolution of cooperation
Social dilemmas

ABSTRACT

Behavior in social dilemmas is often inconsistent with the predictions of classical game theory: people (and a wide variety of other organisms) are more cooperative than might be expected. Here we consider behavior in one such social dilemma, the Traveler's Dilemma, that has received considerable attention in the economics literature but is little known among theoretical biologists. The rules of the game are as follows. Two players each choose a value between R and M , where $0 < R < M$. If the players choose the same value, both receive that amount. If the players choose different values v_1 and v_2 , where $v_1 < v_2$, then the player choosing v_1 receives $v_1 + R$ and the player choosing v_2 receives $v_1 - R$. While the players would maximize their payoffs by both choosing the largest allowed value, M , the Nash equilibrium is to choose the smallest allowed value, R . In behavioral experiments, however, people generally choose values much larger than the minimum and the deviation from the expected equilibrium decreases with R . In this paper, we show that the cooperative behavior observed in the Traveler's Dilemma can be explained in an evolutionary framework. We study stochastic evolutionary dynamics in finite populations with varying intensity of selection and varying mutation rate. We derive analytic results showing that strategies choosing high values can be favored when selection is weak. More generally, selection favors strategies that choose high values if R is small (relative to M) and strategies that choose low values if R is large. Finally, we show that a two-parameter model involving the intensity of selection and the mutation rate can quantitatively reproduce data that from a Traveler's Dilemma experiment. These results demonstrate the power of evolutionary game theory for explaining human behavior in contexts that are challenging for standard economic game theory.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The evolution of cooperation is a central topic of interest in the biological and social sciences (Trivers, 1971; Axelrod and Hamilton, 1981; Levin, 2000; Boyd et al., 2003; Panchanathan and Boyd, 2004; Nowak and Sigmund, 2005; Fowler, 2005; Janssen and Bushman, 2008; Helbing and Yu, 2009; Levin, 2009; Wang et al., 2009; Sigmund et al., 2010). Cooperation is found at all levels of the natural world and lies at the heart of modern human societies. Yet cooperation is often costly, creating a social dilemma: cooperating maximizes the group payoff, but individuals do best by being selfish. In the context of evolutionary game theory, the Prisoner's Dilemma (Trivers, 1971; Axelrod and

Hamilton, 1981; Milinski, 1987; Kraines and Kraines, 1989; Fudenberg and Maskin, 1990; Nowak and Sigmund, 1992, 1993, 2005; Lotem et al., 1999; Wedekind and Milinski, 2000; Ohtsuki and Iwasa, 2004, 2006; Brandt et al., 2005; Imhof et al., 2005, 2007; Pacheco et al., 2006; Fu et al., 2007; Worden and Levin, 2007; Helbing and Yu, 2009; Rand et al., 2009) and the public goods game (Levin, 2000; Hauert et al., 2002; Boyd et al., 2003; Panchanathan and Boyd, 2004; Fowler, 2005; Hauert et al., 2007; Janssen and Bushman, 2008; Levin, 2009; Ohtsuki et al., 2009; Wang et al., 2009; Sigmund et al., 2010; Rand and Nowak, 2011) are the standard paradigms for exploring social dilemmas. However, other interesting social dilemmas also exist. In this paper, we study the evolution of strategies in the Traveler's Dilemma (Basu, 1994), a game that has received little attention among evolutionary theorists (an exception is Smead (2008), who that usedevolutionary simulations to study a version of the Centipede Game that is similar to the Traveler's Dilemma).

* Corresponding author at: Department of Mathematics, Harvard University, Cambridge, MA 02138, USA. Tel.: +1 617 496 4737; fax: +1 617 496 4629.

E-mail address: martin_nowak@harvard.edu (M.A. Nowak).

Just as the Prisoner's Dilemma is often characterized by a standard narrative (involving two prisoners charged with a crime), so too is the Traveler's Dilemma. Imagine that two travelers have purchased identical souvenirs while visiting a remote island. Their airline loses both souvenirs as the two travel home, and the airline then asks each (separately and in private) to declare the value of the souvenir to the nearest dollar. If both travelers declare the same value, then the airline will compensate both with that amount. But if the travelers declare different values, the airline will pay both the smaller of the two. Furthermore, it will reward the traveler claiming the smaller value with a bonus of R and penalize the traveler claiming the larger value the same amount R . For instance, if one traveler claims the souvenir was worth 50 while the other claims it was worth 80, and if $R=2$, then the first will receive 52 and the second 48.

For simplicity, we assume that R is an integer, and to make the game a social dilemma we assume that $R \geq 2$. Furthermore, the game is typically structured so that players' payoffs cannot drop below zero (the airline cannot impose fines), so the players cannot declare values less than R . Fig. 1 shows a (partial) payoff matrix for the game: each player's strategy is the value he or she claims for the item.

What should the travelers do? It might seem natural for both to declare the maximum value M . However, this is not rational (assuming that both players seek to maximize their monetary payoffs). One player could switch his claim to $M-1$ and thereby earn a payoff of $M-1+R > M$ (while reducing the other player's payoff to $M-1-R$). Continuing this iterative deletion of dominated strategies, one finds that the Nash equilibrium is to choose the minimum value, claiming the souvenir was worth just R . As in the Prisoner's Dilemma, "defection" (declaring a low value) dominates "cooperation" (declaring a high value) even though the travelers would be better off if both claimed high values. Thus, the Traveler's Dilemma presents a social dilemma involving coordination coupled with the temptation to exploit the other party.

The Traveler's Dilemma can be used to model situations not only among humans but also among other organisms. For example, competitive egg ejection in the Greater Ani, a communally nesting bird species (Riehl and Jara, 2009; Riehl, 2011), can be modeled by the Traveler's Dilemma. In a nest with two females, each female chooses a time to switch from ejecting eggs out of the nest to laying eggs. If both chose an early date (i.e., a large claim), then both successfully lay a large number of eggs and earn a large (fitness) payoff. But if one chooses to wait, she will eject the other's already-laid eggs and achieve an even higher payoff while inflicting a loss on the earlier-laying female.

Payoffs for row player

	97	98	99	100
97	97	99	99	99
98	95	98	100	100
99	95	96	99	101
100	95	96	97	100

Fig. 1. A partial payoff matrix for the Traveler's Dilemma when $R=2$ and the maximum claim $M=100$. Payoffs are shown for the row player (blue). When the travelers claim the same value, they are both awarded that amount (red). But if one traveler claims a value L and the other a value $H > L$, then the former receives $L+R$ and the latter receives $L-R$. For example, if the row player claims 98 and the column player 99, the row player gets $98+R=100$ and the column player $98-R=96$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Human behavior is the main focus of the current paper. When people are asked to play the Traveler's Dilemma in the laboratory, their behavior deviates significantly from the predictions of classical game theory. For example, Capra et al. (1999) conducted a repeated Traveler's Dilemma in which the value of R varied across treatments. Although the Nash equilibrium is to declare the minimum value regardless of R , Capra et al. (1999) found a significant inverse relationship between R and the average claim. This relationship held both in early rounds before any in-game learning occurred and in later rounds after subjects gained experience.

Capra et al. (1999) and Goeree and Holt (1999) proposed learning models based on the logit probabilistic decision rule to explain the development of the inverse relationship between R and the average claim over the course of the experimental session. Goeree and Holt (1999) adapted this rule to model the "introspection" that occurs before the game begins to explain behavior in initial rounds before in-game learning becomes a factor. They found that if the cognitive "noise" (defined in a suitably precise way) increases with each successive round of introspection, they could reproduce (qualitatively) the inverse relationship between R and the average claim.

In this paper, we instead propose an evolutionary model to explain people's behavior in the Traveler's Dilemma. We study stochastic population dynamics in finite populations with varying intensity of selection and varying mutation rate (Nowak et al., 2004). Critically—and unlike logit learning models or "introspection"—our model does not require any notion of cognition or rationality. We show that stochastic evolutionary models can explain the observed behavior in early rounds, reflecting the (genetically or culturally) evolved intuitions subjects bring with them into the laboratory. We also derive analytic results showing how the evolutionarily favored strategies vary as a function of R and the maximum claim M . In our model, higher payoff strategies are more likely to reproduce on average, but random chance also plays an important role. Sometimes lower payoff strategies reproduce, and sometimes higher payoff strategies die out. When the intensity of selection is low in the context of imitation, people are uncertain when evaluating their own payoffs and the payoffs of others.

We also consider the role of mutation: sometimes offspring do not inherit their parent's strategy but instead assume a random new strategy. In the context of imitation dynamics, this means that players sometimes get confused when trying to imitate a higher payoff individual and adopt the wrong strategy. The higher the mutation rate, the greater the uncertainty about the strategies of others.

When selection is weak, the dynamics depend greatly on the mutation rate μ . In the high mutation limit, $\mu \rightarrow 1$, all strategies are present at approximately equal abundances at the same time (Antal et al., 2009; Traulsen et al., 2009). Thus, success is determined not by one's ability to resist invasion but rather by one's performance against a population uniformly distributed over all strategies. Put another way, the optimal strategy is the one that maximizes its expected payoff against a uniform distribution of opposing strategies.

In the low mutation limit, $\mu \rightarrow 0$, novel mutants will either die out or completely take over the population before a new mutant arises (Nowak et al., 2004; Fudenberg and Imhof, 2006; Hauert et al., 2007, 2008). The population makes transitions between homogeneous states in which all agents in the population play the same strategy. Hence, although all strategies are still present at equal frequency in the steady state distribution, at most two strategies are present in the population at the same time. Therefore, success in the low mutation limit is determined by one's ability to resist invasion by a single (randomly chosen) opponent. It is not expected payoff against a uniform distribution of

strategies that determines success but rather expected relative payoff in pairwise competitions with a single random opposing strategy (Antal et al., 2009).

We can summarize the observations above as follows. Consider two strategies, one claiming a and one claiming b , $b < a$. Call the former strategy A and the latter strategy B . In A – A interactions, A receives a payoff of a . In A – B interactions, A receives a payoff of $b - R$. In B – B interactions, B receives a payoff of b . And in A – B interactions, B receives a payoff of $b + R$.

In the high mutation, weak selection limit, A and B are equally abundant in the population. A 's average payoff is thus $a + b - R$ and B 's average payoff is $b + b - R$. When R is small, $a + b - R > b + b + R$, so the more “cooperative” strategy A is favored over the less cooperative B . When R is large, on the other hand, the opposite is true and B is favored over A .

In the low mutation, weak selection limit, A can be favored over B even though B beats A in pairwise interactions ($b - R < b + R$) since A does better against itself than B does against itself. When R is small, a population of A -players can be resistant to invasion by B -players, and A can in fact have a larger basin of attraction than B (when $a + b - R > b + b + R$).

In both cases, strategies that are not Nash equilibria can be favored by selection. Because the population is far from equilibrium, traditional game-theoretic solution concepts have little relevance. Instead, we ask what strategy is the most common in the stationary distribution of our stochastic model (see Section 2). The inclusion of the learning errors represented by weakening selection and increasing mutation are relevant and realistic: many real-world processes are in fact far from equilibrium and learning is rarely perfect. Importantly, our non-equilibrium calculations can explain laboratory behavior that traditional equilibrium approaches have had little success in explaining.

This paper is organized as follows. In Section 2, we introduce the model and describe the evolutionary process we are studying. In Sections 3–5, we derive analytic results showing how the favored and most frequent strategies vary as R changes in the limit of weak selection. In Section 6, we describe how earlier results can be used to determine the optimal choice of R from the perspective of the airline. Given that selection in the real-world can be weak (but not necessarily in the limit of weak selection), we examine the model via computer simulations in the case of arbitrary selection strength and mutation rates in Section 7. In Section 8, we find the values of the selection strength and mutation rate that best fit the data observed by Capra et al. (1999). In Section 9, we discuss our models in relation to other standard ones. Finally, we conclude in Section 10.

2. Evolutionary process

We examine the Traveler’s Dilemma from the perspective of evolutionary game theory (Maynard Smith, 1982; Hofbauer and Sigmund, 1988, 1998; Weibull, 1997; Samuelson, 1998; Cressman, 2003; Nowak and Sigmund, 2004; Imhof and Nowak, 2006; Gintis, 2009; Sigmund, 2010) in finite populations (Nowak et al., 2004; Taylor et al., 2004).

Each player chooses a strategy from a discrete space consisting of $n = M - R + 1$ strategies (the value she is declaring between R and M). We index these strategies R, \dots, M . The payoff matrix $A = (a_{ij})$, $R \leq i, j \leq M$, where the entry a_{ij} is the payoff of a traveler declaring a value i when the other traveler declares a value j , is given by

$$a_{ij} = \begin{cases} i & \text{if } i = j, \\ i + R & \text{if } i < j, \\ j - R & \text{if } i > j. \end{cases} \quad (1)$$

Suppose that there are N individuals in the population and N_k of them are playing strategy k ($N = N_R + \dots + N_M$). We assume that the population is well-mixed and that players interact randomly. Then the expected payoff (up to a constant factor) of an individual playing strategy i is

$$\pi_i = \sum_{k=R}^M N_k a_{ik} - a_{ii}. \quad (2)$$

The fitness f_i of an individual with strategy i is given by $f_i = \exp(\delta \pi_i)$ (Traulsen et al., 2008). While the payoff π_i measures how well an individual is performing in the game, one’s reproductive fitness may depend on factors other than π_i . One could be engaged in multiple distinct games, for example, with the game in question only making a small contribution to one’s fecundity. The parameter δ , which we call the selection intensity, determines the nature of the payoff-to-fitness mapping. In the limit $\delta \rightarrow 0$, all individuals have roughly the same reproductive success. In the limit $\delta \rightarrow \infty$, even small payoff differences lead to large differences in reproductive success: a small payoff advantage results in a disproportionately large reproductive advantage.

Evolution occurs via the frequency-dependent Moran process. In each time step, an individual is chosen at random to be a “child” (or “learner”). That individual copies, with probability $1 - \mu$, the strategy of a “parent” (or “teacher”) chosen from the population with probability proportional to its fitness. With probability μ , the learner adopts one of the n strategies at random. Because mutation is possible from any one strategy to any other strategy, the random process describing the state of the population is ergodic. In particular, this means that it has a unique stationary distribution.

3. Low mutation, weak selection limit

In the absence of any payoff differences (or when $\delta = 0$), each of the n strategies has frequency $1/n$ in the stationary distribution. Thus, when payoffs are different and individuals are subject to selection, we say that selection favors a strategy if its frequency in the stationary distribution is $> 1/n$. We say that selection opposes a strategy if its frequency in the stationary distribution is $< 1/n$.

When selection is weak, each strategy’s frequency in the stationary distribution is still close to $1/n$ —there is just a small perturbation from the neutral value. Antal et al. (2009) showed that when $\delta \rightarrow 0$ and $\mu \ll 1/N$, the perturbation for strategy k is proportional to

$$L_k = \frac{1}{n} \sum_{i=R}^M (a_{kk} + a_{ki} - a_{ik} - a_{ii}). \quad (3)$$

Hence, strategy k is favored by selection when $L_k > 0$. We can interpret this condition as follows. When $a_{kk} + a_{ki} > a_{ii} + a_{ik}$, a k -mutant in a population of i -players has a higher probability of taking over the population than does an i -mutant in a population of k -players. The condition (3) thus says that strategy k on average can invade and take over uniform populations of other strategies. Conversely, a uniform population of k -players is on average resistant to invasion by other strategies.

Using the payoff matrix (1), we find that

$$L_k = \frac{(M - 5R + 1)(2k - M - R)}{2(M - R + 1)}. \quad (4)$$

L_k is an increasing function of k when $R < (M+1)/5$ and a decreasing function of k when $R > (M+1)/5$. Thus,

$$L_k > 0 \iff \begin{cases} R < \frac{M+1}{5} & \text{and } k \in \left(\frac{M+R}{2}, M\right], \\ R > \frac{M+1}{5} & \text{and } k \in \left[R, \frac{M+R}{2}\right). \end{cases} \quad (5)$$

When R is large, strategies claiming larger values are favored. When R is small, strategies claiming smaller values are favored. When $R = (M+1)/5$, then (weak) selection has no effect and each strategy has the same frequency $1/n$.

The strategy k that maximizes L_k is the one that is most frequent in the stationary distribution. Since L_k is either monotonically increasing or decreasing, it follows that

$$L_k \text{ is maximized} \iff \begin{cases} R < \frac{M+1}{5} & \text{and } k = M, \\ R > \frac{M+1}{5} & \text{and } k = R. \end{cases} \quad (6)$$

When R is large relative to the maximum claim M , there is a strong incentive for individuals to declare low values as the reward for declaring the lower of the two values is large. Thus, the most successful strategy makes the minimum claim R . When the reward is small relative to M , on the other hand, the high payoffs obtained when both players claim large values trump the reward to be had by declaring a low value. Thus, the most successful strategy makes the maximum claim M .

For $M=100$ and $R=2$, L_k is maximized for $k=100$. The strategy making the highest possible claim is thus chosen when the selection intensity and mutation rate are both small.

4. High mutation

In the high mutation case, $\mu \gg 1/N$, strategy k is favored by selection if (Antal et al., 2009)

$$H_k = \frac{1}{n^2} \sum_{i=R}^M \sum_{j=R}^M (a_{kj} - a_{ij}) > 0. \quad (7)$$

To understand this condition, first observe that when mutation is high, all strategies are (roughly) equally abundant. Recall that since δ is small, $f_i = \exp(\delta\pi) \approx 1 + \delta\pi$. Hence, the fitness of strategy k is $f_k = 1 + (\delta/n) \sum_{i=1}^n a_{ki}$. By comparing this to the average fitness $(1/n) \sum_{j=1}^n f_j$, we obtain (7).

Using the payoff matrix (1), we find that

$$H_k = -\frac{k^2}{2(M-R+1)} + \frac{k(1+2M-4R)}{2(M-R+1)} + \frac{-2M^2 + M(4R-1) + R(7R-2)}{6(M-R+1)}. \quad (8)$$

In the extreme cases when $k=R$ (the lowest possible valuation) or $k=M$ (the highest), we have

$$H_R = -\frac{(M-R)(2M-8R+1)}{6(M-R+1)} \quad (9)$$

and

$$H_M = \frac{(M-R)(M-7R+2)}{6(M-R+1)}. \quad (10)$$

Thus, the lowest claim (the Nash equilibrium) is favored when $R > (2M+1)/8$ and the largest claim is favored when $R < (M+2)/7$. Note that we cannot have both $H_R > 0$ and $H_M > 0$

since $(2M+1)/8 > (M+2)/7$ for $M \geq 2$. It follows that

$$H_k > 0 \iff \begin{cases} R > \frac{2M+1}{8} & \text{and } k \in \left[R, \frac{6M-12R+3+A_{M,R}}{6}\right), \\ \frac{M+2}{7} \leq R \leq \frac{2M+1}{8} & \text{and } k \in \left(\frac{6M-12R+3-A_{M,R}}{6}, \frac{6M-12R+3+A_{M,R}}{6}\right), \\ R < \frac{M+2}{7} & \text{and } k \in \left(\frac{6M-12R+3-A_{M,R}}{6}, M\right], \end{cases} \quad (11)$$

The endpoints of the second interval in (11) are precisely the roots of the quadratic (in k) H_k .

The most frequent strategy when mutation is high is $k = (2M-4R-1)/2$. This is the k that maximizes H_k . To see this directly, we argue as follows. When selection is weak and mutation is high, all strategies are (about) equally abundant. The average payoff of strategy k is

$$\begin{aligned} \bar{\pi}_k &= \sum_{j=R}^{k-1} \pi_{kj} + \pi_{kk} + \sum_{j=k+1}^M \pi_{kj} \\ &= \sum_{j=R}^{k-1} (j-R) + k + \sum_{j=k+1}^M (k+R) \\ &= k + \frac{1}{2}(k-R)(k-R-1) + (k+R)(M-k). \end{aligned} \quad (12)$$

We have

$$\frac{d\bar{\pi}_k}{dk} = \frac{1}{2} + M - 2R - k, \quad (13)$$

so $d\bar{\pi}_k/dk = 0$ when $k = M - 2R + 1/2$. Since strategies are restricted to lie in the interval $R \leq k \leq M$ and $R > 1$, we find that

$$H_k \text{ is maximized} \iff \begin{cases} R \leq \frac{2M+1}{6} & \text{and } k = \left\lceil \frac{2M-4R+1}{2} \right\rceil, \\ R > \frac{2M+1}{6} & \text{and } k = R. \end{cases} \quad (14)$$

In the first case, H_k increases when $k < (2M-4R+1)/2$ and decreases afterwards. In the second case, H_k is a decreasing function of k . As in the low mutation case, the size of the reward R is the determining factor. For $M=100$ and $R=2$, H_k is maximized for $k=96,97$ (since we require strategies to be integer-valued). A strategy is favored by selection, $H_k > 0$, when $k \geq 43$.

5. Any mutation rate, weak selection limit

For an arbitrary mutation rate μ , selection favors strategy k if (Antal et al., 2009)

$$L_k + N\mu H_k > 0 \quad (15)$$

We can determine which strategy is most abundant in the stationary distribution by using the fact that strategy k is more abundant than strategy j if

$$L_k + N\mu H_k > L_j + N\mu H_j. \quad (16)$$

Fig. 2 shows which strategy is most abundant as R and $N\mu$ vary. For R equal to 2, 5, or 10, L_k increases with k . When $N\mu$ is small, L_k dominates $N\mu H_k$ and so $k=100$ is most frequent. When $N\mu$ is large, on the other hand, $N\mu H_k$ dominates L_k and so $k = \lceil (2M-4R+1)/2 \rceil$ is most frequent. This means that high mutation results in a smaller k being the most frequent strategy. In contrast, for R equal to 25, L_k decreases with k . Thus, L_k is maximized for $k=R$ whereas H_k is still maximized for $k = \lceil (2M-4R+1)/2 \rceil$. As $N\mu$ increases, the most frequent strategy claims a larger and larger value.

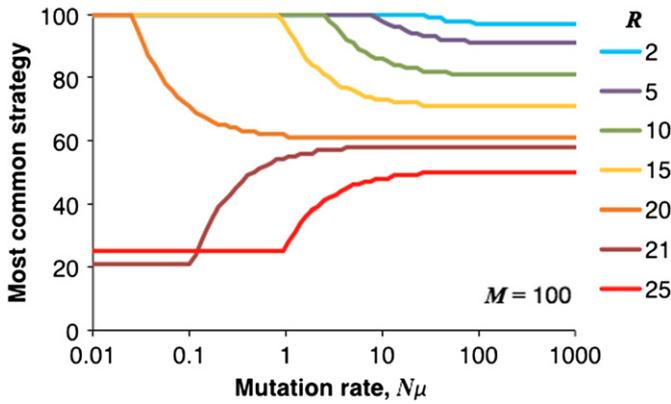


Fig. 2. The most common strategy versus the mutation rate (in terms of the expected number of mutants per generation, $N\mu$). N is the total population size and μ is the mutation rate. The most common strategy is the one that maximizes the quantity (16). Values must be between R and M . When R is small, “cooperative” strategies are most common. When R is large, the opposite is true.

6. Reward size and strategy selection

Now we consider the effect that R has on the average claim. When there is no selection, each of the $n = M - R + 1$ strategies has frequency $1/n$. In the limit of weak selection, the frequency of strategy k is $1/n + \delta_k$, where δ_k is a small perturbation (positive or negative) from neutrality. Antal et al. (2009) showed that δ_k is proportional to L_k in the low mutation case, H_k in the high mutation case, and $L_k + N\mu H_k$ generally, where N is the population size and μ the mutation rate.

When there is no selection, the average claim is $\sum_{k=R}^M (k/n)$. We would like to know the minimum value of R such that the average claim becomes more “cooperative” when passing from the neutral case to the weak selection case. In such a situation, selection can be said to favor higher claims as adding selection to the neutral process increases the average offer.

For this to happen, we must have

$$\sum_{k=R}^M \left(\frac{k}{n} + k\delta_k \right) - \sum_{k=R}^M \frac{k}{n} > 0, \tag{17}$$

or

$$\sum_{k=R}^M k\delta_k > 0. \tag{18}$$

We determine when this inequality holds in the low mutation case. Since L_k is proportional to δ_k , this condition can be written as

$$0 < \sum_{k=R}^M kL_k = \frac{1}{12}(M-R)(M-5R+1)(M-R+2). \tag{19}$$

Since $M > R$, this condition is equivalent to

$$R < \frac{M+1}{5}, \tag{20}$$

the same condition for L_k to be an increasing function of k . When $M=100$, $(M+1)/5 = 20.2$. Thus, when $M=100$ and $R \leq 20$, the average strategy becomes more “cooperative” when we pass from neutrality to weak selection. When $R \geq 21$, the average strategy becomes less “cooperative.” To explore this condition, we compute the steady-state distribution numerically for various intensities of selection without resorting to simulations (Hauert et al., 2007). Fig. 3(a) shows that the average strategy as R and the selection strength vary. The population size is 200. Consistent with the calculated threshold, the average strategy increases with

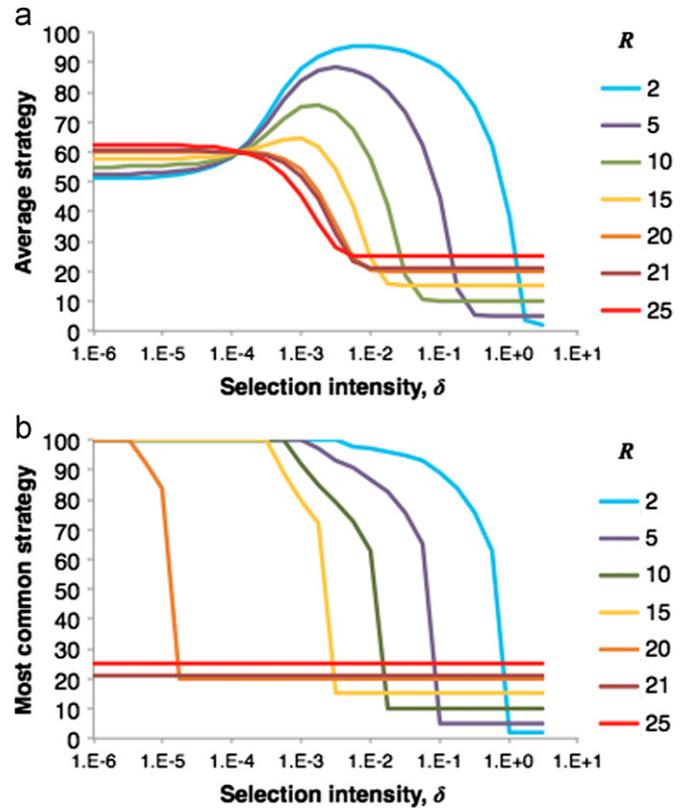


Fig. 3. The average strategy (a) and the most common strategy (b) in the low mutation case as the selection strength and R vary, determined numerically (Hauert et al., 2007) with a population size of 200. (a) When $R < (M+1)/5 = 20.2$, the average strategy initially becomes more cooperative as the selection strength δ increases. As δ gets larger and larger, however, the average strategy approaches the Nash equilibrium of R . When $R > (M+1)/5 = 20.2$, the average strategy decreases monotonically with R . (b) For small values of R , the most common strategy when selection is weak is the cooperative one of claiming the largest possible value, here $M=100$. As δ increases, playing the Nash equilibrium eventually becomes most common. For large values of R , the Nash equilibrium is the most common strategy regardless of selection strength.

the selection strength for $R=2, 5, \text{ and } 10$ whereas the average strategy decreases with selection strength for $R=25$. Thus, if the airline wants to minimize its payments, it should choose an R that is at least 21. However, because the minimum claim increases with R , it is in the interest of the airline to choose an R that is not too much larger than the critical value of $(M+1)/5$.

Fig. 3(b) is the analogous plot for the most frequent strategy. When $\delta \rightarrow 0$ and $R=2, 5, 10$, $k=100=M$ is most frequent. But as the selection strength increases, the most common claims get smaller and smaller, ultimately reaching the Nash equilibrium $k=R$ in the limit $\delta \rightarrow \infty$. On the other hand, when $R=25$, there is a strong-enough incentive to make low claims, so $k=R$ is most frequent regardless of selection strength.

Fig. 4 shows the distribution of strategy frequencies when $R=2$ and $M=100$ as the selection intensity varies. The population size is 200. Again we exactly compute the steady-state distribution numerically without using simulations. We find that when selection is weak, the distribution is concentrated around $k=M=100$. As the selection strength increases, the distribution moves to the left. Players are making smaller claims. Eventually, when selection is very strong, everyone is essentially playing the Nash equilibrium $k=R=2$.

These results have interesting implications for economic mechanism design. As we saw earlier, when selection is weak, the average strategy varies inversely with R . The larger R is, the

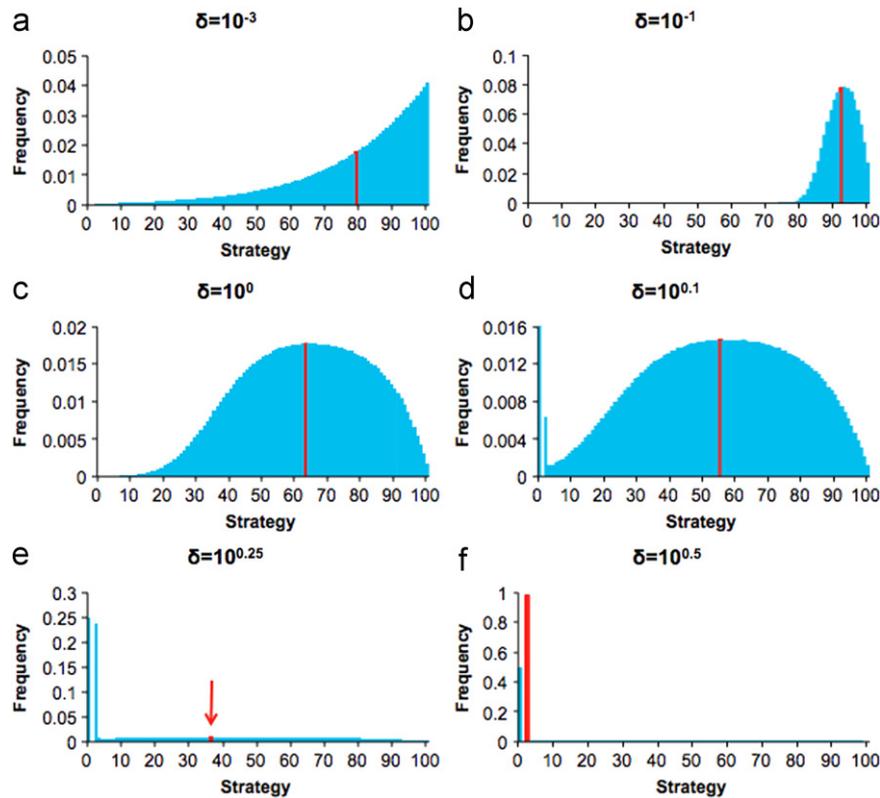


Fig. 4. The strategy distributions in the low mutation case as the selection strength varies, determined numerically with a population size of 200. Here $R=2$ and $M=100$. As the selection strength δ increases, the distribution moves from one with the weight around the “cooperative” strategies to one with the weight around the “selfish” ones. The average strategy for each distribution is marked by a red line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

smaller the typical claims. When selection is strong, on the other hand, everyone claims R (the equilibrium strategy), so the larger R is, the larger the typical claims. If we interpret the selection strength as a measure of how precise people’s information is when learning, then this shows that the choice of the airline’s mechanism (i.e., its choice of R) for reducing the amount it pays in compensation depends critically on assumptions about people’s level of precision. Indeed, past experimental work on the Traveler’s Dilemma suggests that people do not behave rationally: their claims decrease as R increases, consistent with our theory in the weak selection case (Capra et al., 1999).

7. Any mutation rate and selection intensity

We have seen above that we can find the stationary distribution analytically when (1) selection is weak and mutation is low, or (2) mutation is high. We can calculate the distribution numerically in the low mutation case for arbitrary selection intensities. But when we would like to determine the stationary distribution for selection intensity δ and mutation rate μ outside these ranges, we must resort to agent-based simulations.

We performed agent-based simulations to determine how the average strategy varies as the mutation rate μ and the selection intensity δ range over arbitrary intervals. In Fig. 5, $R=2$, $M=100$, and the population size $N=200$. When δ is small, the average strategy is close to $(M-R)/2$, its neutral value. As δ increases, the average strategy increases, similar to the behavior observed in Fig. 3. When R is small, as it is here, the average strategy becomes more cooperative when selection is weak. Finally, as δ gets large, the average strategy declines toward the Nash equilibrium $k=2$.

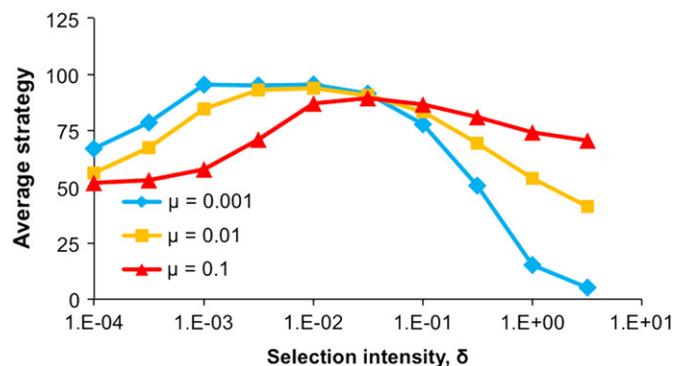


Fig. 5. The average strategy as the selection strength and the mutation rate vary, as determined by agent-based simulations. The population size $N=200$. Results are averaged over 20 simulation runs, each run consisting of 10 million generations (with the stationary values calculated over the last 50% of generations).

8. Comparison with experimental data

Fig. 6 shows the average strategy as a function of R over the first two rounds in Capra et al. (1999)’s experiment. Players were required to make a claim between 80 and 200 (cents) so that the range of possible claims was the same over different treatments (as R varied). Clearly, the average strategy decreases as R increases. This is intuitive as there is more incentive for players to reduce their claims (and approach the Nash equilibrium of 80) when R is large, but it is not consistent with the predictions of classical game theory.

To determine if an evolutionary model could produce the observed behavior, we ran agent based simulations with a population size of $N=200$, varying mutation rates μ (between

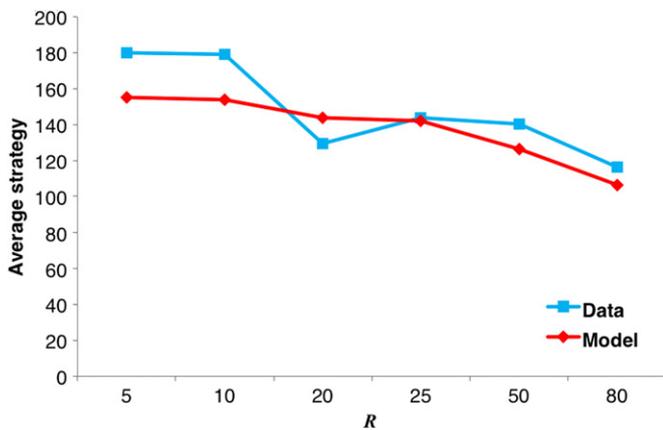


Fig. 6. Data from the first two rounds of a repeated traveler’s dilemma experiment (Capra et al., 1999) and the model of best fit (in the least-squares sense), obtained with agent-based simulations with a selection strength $\delta = 10^{-4}$ and a mutation rate $\mu = 10^{-2.75}$.

10^{-4} and 10^{-1}), and varying intensities of selection δ (between 10^{-5} and 10^{-2}). For each choice of μ and δ , 20 simulations were run, each simulation consisting of 10 million strategy update rounds. The stationary distribution was computed by averaging the frequency of each strategy over the last 5 million rounds. Finally, we obtained average distributions by averaging over all 20 simulations. We then asked which set of parameters minimized the sum of squared differences between the strategy frequencies in the simulation average distribution and those observed in the experimental data. We found that a selection strength of $\delta = 10^{-4}$ and a mutation rate of $\mu = 10^{-2.75}$ resulted in the best fit, parameters that are in the weak selection, low mutation regime. Note that here we are modeling the evolved intuitions that people bring into the lab—and that determine how they play in early rounds—and not the in-game learning that occurs.

The notion of the model-predicted average claim requires some elucidation. It is not necessarily the case that the distribution of strategies in the population eventually becomes fixed. It is possible, for example, that the population continually cycles through various states. What we compute above, both analytically and through agent-based simulations, are the averages in the stationary distribution, which roughly correspond to averages both over the population and over time. In an experiment with 10 subjects, for example, the ten observed claims can be viewed as the results of independent two-step processes: first, the choice of a population, and second, a choice from the stationary distribution of that population. The claims obtained from these choices are independent and identically distributed, the distribution being the stationary one we computed both analytically and through simulations.

9. Discussion

To illustrate the differences between our stochastic evolutionary model and other models, in this section we consider an instance of the Traveler’s Dilemma with a limited strategy space: we take $R=2$ and $M=10$, so there are nine possible strategies (or valuations), $k=2, \dots, 10$. We can summarize the conclusions of various models as follows.

Classical Nash equilibrium: The Nash equilibrium calls for individuals to make the smallest claim possible, $k=2$.

Deterministic evolutionary dynamics: the replicator equation: The replicator equation (Hofbauer and Sigmund, 1998) can be

used to model evolutionary dynamics in the limit of infinite population size (and no mutation). The frequency x_k of strategy k evolves according to the differential equation

$$\dot{x}_k = x_k(f_k - \phi), \tag{21}$$

where f_k is the fitness of strategy k ,

$$f_k = \sum_{j=R}^M x_j a_{kj} \tag{22}$$

(a_{kj} an entry of the payoff matrix (1)), and ϕ is the average fitness,

$$\phi = \sum_{k=R}^M x_k f_k. \tag{23}$$

We computed trajectories numerically for each of the nine strategies and found that strategies other than $k=2$ die out while $k=2$ eventually comes to dominate the population. This is not surprising as the equilibrium of replicator dynamics is generally a Nash equilibrium. Indeed, over a wide variety of choices for R and M , it was always the case that the replicator dynamics converged to the $k=2$ corner of the $(M-R)$ -simplex.

Stochastic evolutionary dynamics: low mutation, weak selection: We use the results of Section 3. Since

$$\frac{M+1}{5} = \frac{11}{5} > 2 = R, \tag{24}$$

L_k is maximized when $k=10$. Furthermore, the strategies favored by selection—those with $L_k > 0$ —are $k=7, \dots, 10$. In contrast to replicator dynamics, the most frequent strategy in this case is $k=10$ —the most “cooperative”—not $k=2$, the least “cooperative.”

Stochastic evolutionary dynamics: high mutation: We use the results of Section 4. Since

$$\frac{2M+1}{6} = \frac{21}{6} > 2 = R, \tag{25}$$

H_k is maximized for

$$k = \left\lceil \frac{2M-4R+1}{2} \right\rceil = \left\lceil \frac{13}{2} \right\rceil, \tag{26}$$

i.e., $k=6$ and $k=7$ are the most abundant strategies in the stationary distribution. This differs from both the low mutation case and the deterministic evolutionary dynamics case.

Logit learning: Capra et al. (1999) conducted a multi-round Traveler’s Dilemma experiment and observed that the average claim decreased as the experiment progressed. They used a logit learning model to explain this decrease and to justify the inverse relationship between R and the average claim at equilibrium.

The logit learning model works as follows. At time t , player p has a set of beliefs about the strategies of other players. These are most easily represented as “counts”: $w_p(j,t)$ represents p ’s belief about the likelihood that someone else is playing a strategy j at time t . A uniform prior is assumed, so $w_p(j,0) = \alpha$ for all j and some fixed α . If player p observes someone playing strategy k at time t , he updates his beliefs by setting $w_p(k,t+1) = w_p(k,t) + \rho$, where the larger ρ is, the more weight p places on recent observations; $w_p(j,t+1) = w_p(j,t)$ for all $j \neq k$.

At time t , p believes that the probability that another individual will play strategy i is

$$\frac{w_p(i,t)}{\sum_j w_p(j,t)}. \tag{27}$$

Given this probability distribution, p can then compute the expected payoff $\pi_p(j,t)$ for each strategy j that he might play. The choice of which strategy to play is then made probabilistically, with the probability that p plays j given by

$$\frac{\exp(\pi_p(j,t)/\mu)}{\sum_k \exp(\pi_p(k,t)/\mu)}. \tag{28}$$

The parameter μ (distinct from our mutation rate μ) represents “cognitive noise.” The smaller it is, the more likely p is to play the strategy with the highest expected payoff. In the limit $\mu \rightarrow \infty$, on the other hand, p picks a strategy to play uniformly at random. As players “learn” according to this rule, an inverse relationship between R and the average strategy develops. To see why this is the case, note that when $\mu > 0$, players make errors in implementing their strategies (they sometimes fail to choose the strategy with the highest expected payoff). When R is small, the penalty for the error is not very large, and the average claim can creep upward. When R is large, on the other hand, errors are penalized severely, and the average strategy remains low.

Capra et al. (1999) fit their experimental data to the logit learning model by finding the ρ and μ that best matched players’ behavior over time. Their goal was to explain the in-game learning that occurs and the inverse relationship between R and the average claim in later rounds. While the logit rule is somewhat similar to an evolutionary process (with exponential fitnesses), it does not explain behavior in early rounds before any in-game learning has occurred. The stochastic evolutionary model we present in this paper, on the other hand, is concerned with explaining the evolved intuitions humans take with them into the lab. Our model justifies the inverse relationship between R and the average claim that Capra et al. (1999) observed in early rounds (Fig. 6) without requiring the abilities of observation (needed in the logit model as players must update their beliefs) or reason.

10. Conclusion

The Nash equilibrium, and generalizations of it, have been the foremost solution concepts in game theory for almost the entire history of the field. However, deviation from equilibrium play has been observed in countless human subjects and in a multitude of games. In the Traveler’s Dilemma, individuals usually make large claims and rarely play the Nash equilibrium of claiming the smallest possible value, R , at least before any in-game learning has occurred. This “irrational” behavior is observed even amongst economists (Becker et al., 2005), the individuals one would think are most likely to analyze the game and play the Nash equilibrium.

Here we have shown how a stochastic evolutionary model can explain why human intuition has evolved to favor such play. There is a population of individuals playing the game, and the evolution of this population is described by a frequency-dependent Moran process. When selection is weak, stochasticity results in higher claims being favored by selection. These results demonstrate the power of stochastic evolutionary game theory for explaining cooperative behavior in contexts outside of the Prisoner’s Dilemma (see also Rand and Nowak, 2012). In a world in which both genetic inheritance and social learning are imperfect, then, “irrational” play can be evolutionarily advantageous, explaining the “anomalous behavior” so often observed by humans playing the Traveler’s Dilemma.

References

- Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C.E., Nowak, M.A., 2009. Mutation-selection equilibrium in games with multiple strategies. *J. Theor. Biol.* 258, 614–622.
- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- Basu, K., 1994. The traveler’s dilemma: paradoxes of rationality in game theory. *Am. Econ. Rev.* 84, 391–395.
- Becker, T., Carter, M., Naeve, J., 2005. Experts playing the traveler’s dilemma. *Hohenheimer Diskussionbeiträge* 252.
- Boyd, R., Gintis, H., Bowles, S., Richerson, P.J., 2003. The evolution of altruistic punishment. *Proc. Nat. Acad. Sci. USA* 100, 3531–3535.
- Brandt, H., Hauert, C., Sigmund, K., 2005. Punishing and abstaining for public goods. *Proc. Nat. Acad. Sci. USA* 103, 495–497.
- Capra, C.M., Goeree, J.K., Gomez, R., Holt, C.A., 1999. Anomalous behavior in a traveler’s dilemma? *Am. Econ. Rev.* 89, 678–690.
- Cressman, R., 2003. *Evolutionary Dynamics and Extensive Form Games*. The MIT Press.
- Fowler, J.H., 2005. Altruistic punishment and the origin of cooperation. *Proc. Nat. Acad. Sci. USA* 102, 7047–7049.
- Fu, F., Chen, X., Liu, L., Wang, L., 2007. Social dilemmas in an online social network: the structure and evolution of cooperation. *Phys. Lett. A* 371, 58–64.
- Fudenberg, D., Imhof, L., 2006. Imitation processes with small mutations. *J. Econ. Theory* 131, 251–262.
- Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. *Am. Econ. Rev.* 80, 274–279.
- Gintis, H., 2009. *Game Theory Evolving*. Princeton University Press.
- Goeree, J.K., Holt, 1999. Stochastic game theory: for playing games, not just for doing theory. *Proc. Nat. Acad. Sci. USA* 96, 10564–10567.
- Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002. Volunteering as red queen mechanism for cooperation in public goods games. *Science* 296, 1129–1132.
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K., 2007. Via freedom to coercion: the emergence of costly punishment. *Science* 316, 1905–1907.
- Hauert, C., Traulsen, A., de Silva, H., Nowak, M.A., Sigmund, K., 2008. Public goods with punishment and abstaining in finite and infinite populations. *Biol. Theor.* 3, 114–122.
- Helbing, D., Yu, W., 2009. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc. Nat. Acad. Sci. USA* 106, 3680–3685.
- Hofbauer, J., Sigmund, K., 1988. *The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection*. Cambridge University Press.
- Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. *Proc. Nat. Acad. Sci. USA* 102, 10797–10800.
- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2007. Tit-for-tat or win-stay, lose-shift? *J. Theor. Biol.* 247, 574–580.
- Imhof, L.A., Nowak, M.A., 2006. Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* 52, 667–681.
- Janssen, M.A., Bushman, C., 2008. Evolution of cooperation and altruistic punishment when retaliation is possible. *J. Theor. Biol.* 254, 541–545.
- Kraines, D., Kraines, V., 1989. Pavlov and the prisoner’s dilemma. *Theor. Decis.* 26, 47–79.
- Levin, S.A. (Ed.), 2000. *Fragile Dominion*. Basic Books.
- Levin, S.A. (Ed.), 2009. *Games, Groups, and the Global Good*. Springer.
- Lotem, A., Fishman, M.A., Stone, L., 1999. Evolution of cooperation between individuals. *Nature* 400, 226–227.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Milinski, M., 1987. Tit for tat in sticklebacks and the evolution of cooperation. *Nature* 325, 433–435.
- Nowak, M., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game. *Nature* 364, 56–58.
- Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428, 646–650.
- Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous populations. *Nature* 355, 250–253.
- Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. *Science* 303, 793–799.
- Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. *Nature* 437, 1291–1298.
- Ohtsuki, H., Iwasa, Y., 2004. How should we define goodness? Reputation dynamics in indirect reciprocity. *J. Theor. Biol.* 231, 107–120.
- Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* 239, 435–444.
- Ohtsuki, H., Iwasa, Y., Nowak, M.A., 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature* 457, 79–82.
- Pacheco, J.M., Santos, F.C., Chalub, F.A.C.C., 2006. Stern-judging: a simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comp. Biol.* 2, e178.
- Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–502.
- Rand, D.G., Nowak, M.A., 2011. The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* 2, 434.
- Rand, D.G., Nowak, M.A., 2012. Evolutionary dynamics in finite populations can explain the full range of cooperative behaviors observed in the centipede game. *J. Theor. Biol.* 300, 212–221.
- Rand, D.G., Ohtsuki, H., Nowak, M.A., 2009. Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *J. Theor. Biol.* 256, 45–57.
- Riehl, C., 2011. Living with strangers: direct benefits favour non-kin cooperation in a communally nesting bird. *Proc. R. Soc. B* 278, 1728–1735.
- Riehl, C., Jara, L., 2009. Natural history and reproductive biology of the communally breeding Greater (Ani) Crotophaga major at gatún lake panama. *Wilson J. Ornithol.* 121, 679–687.
- Samuelson, L., 1998. *Evolutionary games and equilibrium selection*. The MIT Press.
- Sigmund, K., 2010. *The Calculus of Selfishness*. Princeton University Press.
- Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. *Nature* 466, 861–863.

- Smead, R., 2008. The evolution of cooperation in the centipede game with finite populations. *Philos. Sci.* 75, 157–177.
- Taylor, C., Fudenberg, D., Sasaki, A., Nowak, M.A., 2004. Evolutionary game dynamics in finite populations. *B. Math. Biol.* 66, 1621–1644.
- Traulsen, A., Hauert, C., de Silva, H., Nowak, M.A., Sigmund, K., 2009. Exploration dynamics in evolutionary games. *Proc. Nat. Acad. Sci. USA* 106, 709–712.
- Traulsen, A., Shores, N., Nowak, M.A., 2008. Analytic results for individual and group selection of any intensity. *B. Math. Biol.* 70, 1410–1424.
- Trivers, R.L., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Wang, J., Fu, F., Wang, L., 2009. Emergence of social cooperation in threshold public goods games with collective risk. *Phys. Rev. E* 80, 036101.
- Wedekind, C., Milinski, M., 2000. Cooperation through image scoring in humans. *Science* 288, 850–852.
- Weibull, J.W., 1997. *Evolutionary Game Theory*. The MIT Press.
- Worden, L., Levin, S.A., 2007. Evolutionary escape from the prisoner's dilemma. *J. Theor. Biol.* 245, 411–422.