PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Pal S, Lambert M, Nowak MA. 2025 Stabilizing unconditional cooperation. *Proc. R. Soc. A* **481**: 20240945.
https://doi.org/10.1098/rspa.2024.0945

Received: 9 December 2024
Accepted: 28 May 2025

Subject Areas:
computational biology, applied mathematics, mathematical modelling

Keywords:
evolutionary game theory, evolution of cooperation, direct reciprocity

Author for correspondence:
Saptarshi Pal
e-mail: spal@math.harvard.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7879882.

THE ROYAL SOCIETY
PUBLISHING

Stabilizing unconditional cooperation

Saptarshi Pal[1], Mayeul Lambert[3] and Martin A. Nowak[1,2]

[1]Department of Mathematics, and [2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[3]Applied Mathematics Unit, Polytechnic Institute of Paris, France

SP, 0000-0003-3805-1295

The ultimate goal of research on the evolution of cooperation could be conceived as finding a method for stabilizing strategies that always cooperate, that never deviate from cooperation, that never exploit, and never retaliate, because all such activities are inherently problematic. The hope of achieving this cooperative utopia seems unjustified, especially in the context of direct reciprocity, which relies on the principle that cooperative partner strategies succeed in defending themselves by using moderate retaliation. Here, we propose a dynamic that goes a certain way towards achieving the desirable goal. In the ground state, the population consists of individuals that use always-cooperate (ALLC). Occasionally defectors, in the form of always-defect (ALLD), invade the population. Their presence triggers a mutation from ALLC to another strategy, X, with the aim to avert the take-over of defectors. In the absence of X, ALLD dominates ALLC, but in the presence of X, the invasion attempt might fail and subsequently the ALLC ground state can be restored. We study this mutation-selection process in finite and infinite populations. We identify the properties of the ideal rescue strategy. We derive an optimum mutation rate that maximally stabilizes ALLC.

1. Introduction

Cooperation is a central theme in the theory of evolution [1–5]. Cooperators are individuals who pay costs to benefit others. Defectors, by contrast, pay no costs and distribute no benefits. Mechanisms for the evolution of cooperation are interaction structures that allow natural selection to favour cooperation over defection [2,3,6–12]. Here, we focus on the mechanism of direct reciprocity [1,13–19]. There are repeated encounters between the same two individuals. If I cooperate now, I might induce your cooperation later. If I defect now, I might evoke your retaliation. Therefore, in an iterated Prisoner's Dilemma, defection is no longer the only Nash equilibrium [20].

Much of the research within direct reciprocity deals with the question of how to play the repeated game. The search is for strategies that do well. Recently, strategies have been separated into two large categories: rivals and partners [18,21]. Rivals always want a higher payoff than their co-player. They insist on winning. Examples of rivals are always defect (ALLD) or so-called extortionists [22]. Partners accept that, in some situations, their payoff will be lower than that of the co-player, but they incentivize their co-players to cooperate with them: when playing against a partner strategy you maximize your own payoff when you fully cooperate. Partners are focused on sharing. Examples of partners are generous tit-for-tat [23] and win-stay, lose-shift [24]. Broadly speaking, rivals destroy but partners facilitate evolution of cooperation.

When it comes to human behaviour, an inherent problem of direct reciprocity is the need to retaliate against defection. A genuinely good person may be reluctant do to so. But if everyone uses unconditional cooperation then the mechanism of direct reciprocity fails to operate. In this paper, we explore the question of whether it is possible to stabilize a population of always cooperate (ALLC) individuals. We explore the following setting. In the ground state, the population consists only of ALLC players. Occasionally, the ground state is challenged by invasion of defectors. For simplicity but without losing generality, we assume the challengers use the notorious ALLD strategy. When the invasion occurs a 'danger signal' triggers the mutation from ALLC to a rescue strategy, X. The hope is that in the presence of X the invasion attempt will fail and the ground state will be restored.

In the following, we study both stochastic and deterministic evolutionary dynamics of the resulting mutation selection system. First, we study the stochastic evolutionary process by considering a finite population model. In this model, ALLC individuals probabilistically generate a mutant offspring X only when the number of defectors surpasses a certain threshold in the population. Our goal is to compute the likelihood that the ALLC population is restored once a single ALLD mutant attempts to invade it. We examine the effect of the conditional mutation probability, μ, on the restoration probability and characterize properties of the best rescue strategy, X. Later, we study deterministic evolutionary dynamics by considering infinite population size. Here, we find an optimal mutation rate, μ_1, for which ALLC achieves maximal frequency in stable coexistence with X and ALLD.

2. The repeated donation game and strategies

(a) Description of the game

The donation game involves two players, each having two possible actions, cooperation, C, or defection, D. When an individual cooperates, they incur a cost c to provide a benefit b to the co-player. We have $b > c > 0$. When an individual defects, they provide no benefit and thereby incur no cost. The donation game is a special case of the Prisoner's Dilemma. For the four possible outcomes of the game, the payoffs for the row player are given by the payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array}. \tag{2.1}$$

In the repeated donation game, players continue with probability δ after each round. We focus on the limit $\delta \to 1$, which represents the infinitely repeated game. We assume that players make rare implementation errors: they sometimes play D when they intend to play C and *vice versa* [15].

(b) Strategies

A strategy is defined by the probability to cooperate in the first round and by the probabilities to cooperate after every sequence of past play. Here, we study strategies that base their decision on the outcome of the previous round, which are the so-called memory-1 strategies [15]. Since we consider the infinitely repeated game with implementation errors, we can ignore the initial move [15,25].

Therefore, a memory-1 strategy is defined by four parameters, $(p_{CC}, p_{CD}, p_{DC}, p_{DD})$, which denote the probabilities to cooperate if the outcome of the previous round was CC, CD, DC or DD, respectively. In this notation, the first letter (C or D) denotes the move of the focal player, while the second letter (C or D) denotes the move of the other player. A strategy is called deterministic if all probabilities are binary variables taking the values 0 or 1. There are 16 deterministic memory-1 strategies. They include ALLD $(0,0,0,0)$, GRIM $(1,0,0,0)$, win-stay, lose-shift $(1,0,0,1)$, TFT $(1,0,1,0)$ and firm-but-fair (FBF) $(1,0,1,1)$, which is also known as Forgiver and ALLC $(1,1,1,1)$.

(c) Computing expected payoffs

The first step in calculating expected payoff per round for the two players is to determine how often the four states—CC, CD, DC and DD—occur based on the strategies that the players adopt. Earlier work [25,26] details the method for computing the stationary distribution of the four states, provided both players employ memory-1 strategies. Due to rare implementation errors, there is a unique stationary distribution, $(v_{CC}, v_{CD}, v_{DC}, v_{DD})$ [25,26]. The expected payoff of the focal player is then given by $\pi_1 = v_{CC}(b-c) + v_{CD}(-c) + v_{DC}(b) + v_{DD}(0)$. We use $\pi(s_i, s_j)$ to denote the payoff of the player, who adopts the strategy s_i against her co-player who adopts s_j.

The payoffs when both players employ deterministic memory-1 strategies are given in table 1. The effect of implementation error is noticeable when one observes the payoff that the strategy TFT (S_{10} in table 1) obtains against itself. In a game with no errors, two TFT players, who start with cooperation, continue to cooperate in every round and receive an expected per-round payoff of $b-c$. However, with errors, one can derive that two TFT players visit each of the four states equally often: their stationary distribution is $(0.25, 0.25, 0.25, 0.25)$. As a result, they receive an expected payoff of $(b-c)/2$. Therefore, in the presence of noise, TFT fails to achieve the payoff for full cooperation. Consequently, generous tit-for-tat (GTFT) $(1,q,1,q)$ for some $q>0$ is a much better strategy, because it has the ability to forgive [15].

3. Evolutionary dynamics in a finite population

Consider a finite, well-mixed population of size N, whose individuals interact in a repeated donation game. We study evolutionary dynamics in a birth–death process [27] with mutation. In the beginning, the population is only composed of individuals who adopt ALLC. We call this the 'ground state' of the population. We challenge the ground state by switching the strategy of one individual to ALLD. We are interested in studying the evolutionary dynamics and the final state of the process that is initiated with this alteration. Below, we describe in detail the evolutionary process that unfolds.

(a) Description of the birth–death process with conditional mutations

At each step of the birth–death process, first an individual is chosen randomly from the population, with probability proportional to its current fitness, to produce an offspring. In the second step, a second individual, selected uniformly at random to die. Thereby, the population remains at a constant size. During the birth event, ALLD individuals reproduce without mutation; they always produce ALLD offspring. However, when an ALLC individual reproduces, the offspring is ALLC with probability $1-\mu_e$, but adopts strategy X with probability μ_e. Individuals

Table 1. Payoffs for the row player when both players employ deterministic memory-1 strategy in an infinitely repeated donation game with implementation errors [25].The $(p_{CC}, p_{CD}, p_{DC}, p_{DD})$ form of the memory-1 strategy S_k is Binary(k) with p_{DD} being the least significant bit.

	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
S_0	$\mathbf{0}$	$b/2$	0	$b/2$	$b/3$	b	$b/2$	b	0	$b/2$	0	$b/2$	$b/2$	b	$2b/3$	b
S_1	$-c/2$	$\mathbf{b - c/2}$	$b - c/3$	$b - c/2$	$b - 2c/5$	$2b - c/3$	b	$3b - c/4$	$-c/2$	$2b - c/3$	$b - c/3$	$2b - c/3$	$2b - c/4$	b	b	b
S_2	0	$b - c/3$	$\mathbf{b - c/4}$	$b - c/2$	0	$2b - c/3$	0	$2b - c/3$	0	$b - c/3$	$b - c/3$	$b - c/2$	$2b - c/4$	$2b - c/2$	$4b - 2c/5$	$2b - c/2$
S_3	$-c/2$	$b - c/2$	$b - c/2$	$\mathbf{b - c/2}$	$-c/2$	$b - c/2$	$b - c/2$	$b - c/2$	$-c/2$	$b - c/2$	$b - c/2$	$b - c/2$	$b - c/2$	$2b - c/2$	$2b - c/2$	$2b - c/2$
S_4	$-c/3$	$2b - c/5$	0	$b/2$	$\mathbf{b - c/4}$	b	$b/3$	b	$-c/3$	$2b - c/5$	0	$b/2$	$3b - c/6$	b	$2b/3$	b
S_5	$-c$	$b - 2c/3$	$b - 2c/3$	$b - c/2$	$-c$	$\mathbf{b - c/2}$	$b - c/2$	$2b - c/3$	$-c$	$b - c/2$	$b - c/2$	$2b - c/3$	$b - c/2$	b	b	b
S_6	$-c/2$	$-c$	0	$b - c/2$	$-c/3$	$b - c/2$	$\mathbf{0}$	$2b - c/3$	$-2c/3$	$-c$	$b - c/2$	$2b - 2c/3$	$b - c/2$	$4b - 3c/5$	$4b - 2c/5$	$2b - c/2$
S_7	$-c$	$b - 3c/4$	$b - 2c/3$	$b - c/2$	$-c$	$b - 2c/3$	$b - 2c/3$	$\mathbf{b - c/2}$	$-c$	$-c$	$2b - 2c/3$	$2b - 2c/3$	$2b - 3c/4$	$4b - 3c/5$	$2b - c/2$	$2b - c/2$
S_8	0	$b/2$	0	$b/2$	$b/3$	b	$2b/3$	b	$\mathbf{0}$	$3b - c/5$	0	$3b - c/5$	$3b - c/6$	$3b - c/3$	$3b - c/4$	$3b - c/3$
S_9	$-c/2$	$b - 2c/3$	$b - c/3$	$b - c/2$	$b - 2c/5$	$b - c/2$	b	b	$b - 3c/5$	$\mathbf{b - c}$	$b - c/2$	$b - c$	$b - c/2$	$3b - 2c/3$	$3b - c/3$	$2b - c/2$
S_{10}	0	$b - c/3$	$b - c/3$	$b - c/2$	0	$b - c/2$	$b - c/2$	$2b - 2c/3$	0	$b - c/2$	$\mathbf{b - c/2}$	$2b - 2c/3$	$b - c/2$	$b - c$	$b - c$	$b - c$
S_{11}	$-c/2$	$b - 2c/3$	$b - c/2$	$b - c/2$	$-c/2$	$b - 2c/3$	$2b - 2c/3$	$2b - 2c/3$	$b - 3c/5$	$b - c$	$2b - 2c/3$	$\mathbf{3b - 3c/4}$	$2b - 3c/4$	$b - c$	$b - c$	$b - c$
S_{12}	$-c/2$	$b - 2c/4$	$b - 2c/4$	$b - c/2$	$b - 3c/6$	$b - c/2$	$b - c/2$	$3b - 2c/4$	$b - 3c/6$	$b - c/2$	$b - c/2$	$3b - 2c/4$	$\mathbf{b - c/2}$	$5b - 3c/6$	$5b - 3c/6$	$2b - c/2$
S_{13}	$-c$	$-c$	$b - 2c/2$	$b - 2c/2$	$-c$	$-c$	$3b - 4c/5$	$3b - 4c/5$	$b - 3c/3$	$2b - 3c/3$	$b - c$	$b - c$	$3b - 5c/6$	$\mathbf{3b - 3c/4}$	$3b - 2c/3$	$3b - 2c/3$
S_{14}	$-2c/3$	$-c$	$2b - 4c/5$	$b - 2c/2$	$-2c/3$	$-c$	$2b - 4c/5$	$b - 2c/2$	$b - 3c/4$	$b - 3c/3$	$b - c$	$b - c$	$3b - 5c/6$	$2b - 3c/3$	$\mathbf{b - c}$	$b - c$
S_{15}	$-c$	$-c$	$b - 2c/2$	$b - 2c/2$	$-c$	$-c$	$b - 2c/2$	$b - 2c/2$	$b - 3c/3$	$b - 2c/2$	$b - c$	$b - c$	$b - 2c/2$	$2b - 3c/3$	$b - c$	$\mathbf{b - c}$

who adopt X reproduce without mutation. We allow X to be any memory-1 strategy. Later, we also consider back mutation from X to ALLC.

At any point in time, the state of the population is specified by three numbers (N_C, N_D, N_X), denoting the abundance of ALLC players, ALLD players and X players. The total population size, $N = N_C + N_D + N_X$, is constant.

The mutation probability from ALLC to X, μ_e depends on the number of ALLD players in the population, N_D. In particular, we consider

$$\mu_e = \begin{cases} \mu & \text{if } N_D \geq k, \\ 0 & \text{if } N_D < k. \end{cases} \tag{3.1}$$

Thus, ALLC produces mutants only if N_D is greater than or equal to k. We consider $k \geq 1$ in our analysis. ALLC does not mutate if $N_D = 0$. This threshold mutation rate represents a specific choice. Later, we consider a mutation rate that is a linear function of N_D.

We illustrate the evolutionary process in figure 1. For $\mu = 0$, the stochastic process reduces to the fixation dynamics of an ALLD mutant in an ALLC population (see figure 1A). In this case, the population is always composed of at most two strategies, ALLC and ALLD, and there are two absorbing states, $(N, 0, 0)$ and $(0, N, 0)$. For $\mu > 0$, the population is composed of at most three strategies, ALLC, ALLD and X. Since mutations only occur when ALLD is present in the population, the population stabilizes once it returns to the ground state, $(N, 0, 0)$. Likewise, since reproductions by ALLD and X are never subject to mutations, the population also stabilizes when it reaches the homogeneous states $(0, N, 0)$ or $(0, 0, N)$. Thus, there are three final states of the stochastic process that begins from $(N-1, 1, 0)$: the ground state, $(N, 0, 0)$, the state where everyone adopts ALLD, $(0, N, 0)$ and the state where everyone adopts X, $(0, 0, N)$ (see figure 1B).

(b) Fitness of strategies in a population

In the birth–death process, individuals are selected for reproduction with probability proportional to their current fitness. We assume that fitness is a positive quantity that monotonically increases with average payoff. We compute the average payoff of a strategy i (denoted Π_i), by weighing its expected per-round payoffs against other strategies, $\pi(i, \cdot)$, with the probability of meeting them in the well-mixed population. For the population state $\mathbf{N} := (N_C, N_D, N_X)$, we obtain the payoff values

$$\left.\begin{aligned} &\Pi_C(\mathbf{N}) = (b-c)\left(\frac{\max\{N_C-1, 0\}}{N-1}\right) - c\left(\frac{N_D}{N-1}\right) + \pi(C, X)\left(\frac{N_X}{N-1}\right) \\ &\Pi_D(\mathbf{N}) = b\left(\frac{N_C}{N-1}\right) + \pi(D, X)\left(\frac{N_X}{N-1}\right) \\ \text{and}\qquad &\Pi_X(\mathbf{N}) = \pi(X, C)\left(\frac{N_C}{N-1}\right) + \pi(X, D)\left(\frac{N_D}{N-1}\right) + \pi(X, X)\left(\frac{\max\{N_X-1, 0\}}{N-1}\right). \end{aligned}\right\} \tag{3.2}$$

We assume that an individual's fitness is an exponential function of its average expected payoff [28]. That is, an individual with strategy i in the population state $\mathbf{N}$ has the fitness $\mathrm{F}_i(\mathbf{N}) = e^{\beta \Pi_i(\mathbf{N})}$. Here β is the intensity of selection. We consider $\beta = 1$ from here onwards.

(c) Computing absorption probability into final states

When mutations are absent ($\mu = 0$), the evolutionary process is a discrete-time Markov chain in the state space

$$\mathrm{S}_{\mu=0} = \{(N_C, N - N_C, 0) \mid N_C \in \{0, 1, .., N\}\}. \tag{3.3}$$

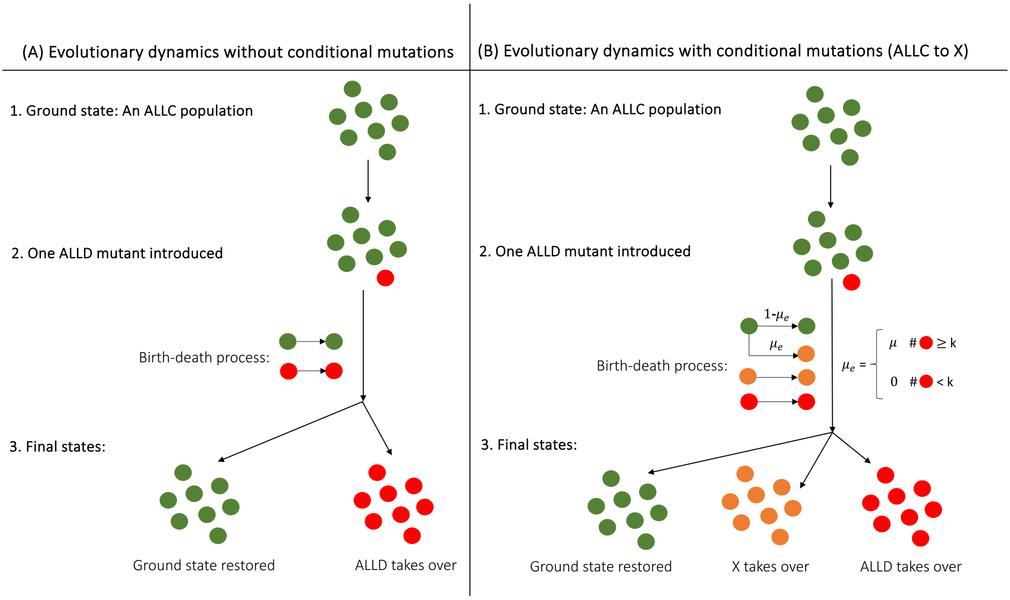


Figure 1. Two evolutionary dynamics involving strategies of the repeated donation game. For both evolutionary dynamics illustrated in this figure, we consider the same starting point. The population that evolves is finite, well-mixed and initially composed of individuals who use the strategy to always cooperate (ALLC) in an infinitely repeated donation game. We call this population the ground state. In this population, a single individual switches to the strategy to always-defect (ALLD). This is the shared starting point of the two distinct dynamics we study from here onwards. (A) *The simple birth–death process*: In this process, first, an individual is randomly selected from the population to give birth, with selection probability proportional to its fitness. Fitness of an individual is defined as the exponential of the expected per-round payoff of its strategy. Next, an individual is selected uniformly at random from the remaining population members to die. The process is repeated from first step. This dynamic results in two possible final states for the population; the initial ground state and the state in which all individuals adopt ALLD. (B) *Birth–death process with condtional mutations*, in an alternate dynamic, we study the birth–death process but with the exception that mutations may happen during the birth event. In particular, with probability μ_e, an ALLC individual gives birth to an offspring that adopts strategy X in the repeated game. With probability $1-\mu_e$ it gives birth to an individual that adopts ALLC. The exact mutation probability depends on the number of ALLD individuals in the current population. In this model, mutations occur only if number of ALLD individuals exceed a certain threshold, k (≥ 1). Reproductions from ALLD and X are not subject to mutations. This dynamic has three final states; the ground state, the state in which everyone adopts X, and the state in which everyone adopts ALLD.

The probability that the process ends in state $(0, N, 0)$, which means all players have adopted ALLD, after beginning from state $(N-1, 1, 0)$ is given by [27]

$$\rho_D = \frac{1}{1+\sum_{k=1}^{N-1}\prod_{j=1}^{k} e^{-c-(b/N-1)}}. \tag{3.4}$$

With the complementary probability, $\rho_C := 1-\rho_D$, the process ends in the other final state, $(N, 0, 0)$ wherein everyone adopts ALLC.

When mutations are possible, $\mu > 0$, the Markov process covers the bigger state space

$$\mathrm{S} := \{(N_C, N_D, N_X) \in \{0, 1, 2, \ldots, N\}^3 \,\Big|\, N_C + N_D + N_X = N\}. \tag{3.5}$$

In addition to $(N, 0, 0)$ and $(0, N, 0)$, the state in which everyone adopts X, $(0, 0, N)$, is also an absorbing state. We define ρ_C, ρ_D and ρ_X as the probabilities that the process finishes in $(N, 0, 0)$,$(0, N, 0)$ and $(0, 0, N)$, respectively, after it begins from state $(N-1, 1, 0)$.

To explicitly compute these absorption probabilities, we first derive the transition matrix **T** associated with the Markov process (see equation (7.1) in **Methods**). Then, we define submatrices **Q** and **R** of **T**. The elements of submatrix **Q** only contain transition probabilities between non-absorbing states in S. It is a square matrix of size $|S| - 3$. The submatrix **R** contains the transition probabilities from non-absorbing states to absorbing states. It is a matrix of size $(|S| - 3) \times 3$. The probability that the process is absorbed in state s_A, when it begins from the non-absorbing state, s_{NA} is given by the element corresponding to the state-pair (s_{NA}, s_A) in the matrix

$$\mathbf{B} := (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}. \tag{3.6}$$

Here, **I** is the identity matrix. The absorption probabilities are thus,

$$\left.\begin{aligned} \rho_C &= B_{((N-1,1,0),(N,0,0))}, \\ \rho_D &= B_{((N-1,1,0),(0,N,0))} \\ \text{and} \qquad \rho_X &= B_{((N-1,1,0),(0,0,N))}. \end{aligned}\right\} \tag{3.7}$$

4. Results

(a) Rescue by deterministic memory-1 strategies

To study whether conditional mutations save a resident ALLC population against an invasion by ALLD, we begin by examining the case where X is a deterministic memory-1 strategy. We first study the dynamics in which mutations are triggered whenever ALLD is present ($k = 1$).

We find that among deterministic memory-1 strategies, TFT provides the best defense against an ALLD invasion (figure 2). While conditional mutations to strategies S_2, S_8, S_9 and S_{14} suppress an ALLD invasion, they do not restore the ground state as effectively as TFT because these rescue strategies themselves out-compete ALLC. Once ALLD is eradicated and mutation stops, the population, which is now composed of X and ALLC, may favour fixation of X over ALLC. For example, this is the case when the rescue strategy X is win-stay-lose-shift [24], S_9. By contrast, when X is TFT then the fixation of ALLC is favoured, because it weakly dominates TFT in the infinitely repeated game, which we consider.

The only other deterministic strategy that displays comparable rescue property is S_{11}. This strategy behaves identically to TFT with the exception that it cooperates if both players defect in the last round—leading to its name *firm-but-fair* [25], FBF. This strategy is also known under the name Forgiver [29]. ALLC also weakly dominates FBF. But since FBF has a higher self-payoff than TFT (see table 1), it has higher fitness than TFT in any mixed population with ALLC. As a result, FBF contributes to a weaker rescue effect than TFT.

For both TFT and FBF, we observe that rescue is most successful for intermediate mutation rates. Although a high mutation rate substantially reduces the chances that ALLD takes over, it risks producing a residual population in which the abundance of X is much greater than that of ALLC. In this case, a return to a homogeneous ALLC ground state is unlikely.

This particular problem disappears if we consider a simple extension of our stochastic process: Once ALLD is eliminated, a back mutation from X to ALLC offspring with probability $\mu_{\text{back}} (> 0)$ is triggered. Now the process is guaranteed to end up in the ground state of only ALLC once ALLD becomes extinct. Interestingly, we observe that the strategy S_2—to cooperate only after DC—provides marginally better rescue than TFT in this process as long as the mutation rate μ is not too high (see electronic supplementary material, figure S3).

We note that the mechanism of conditional mutation backfires against ALLC when the rescue strategy is poorly chosen. For some X, the odds that ALLD takes over increases in comparison to the case where ALLC never mutates ($\mu = 0$). For the specific parameters we adopt in figure 2, this occurs when X is S_5 (anti TFT), S_7 (cooperate always, except after mutual cooperation), S_{13} (cooperate always, except after CD) or S_0 (ALLD itself).

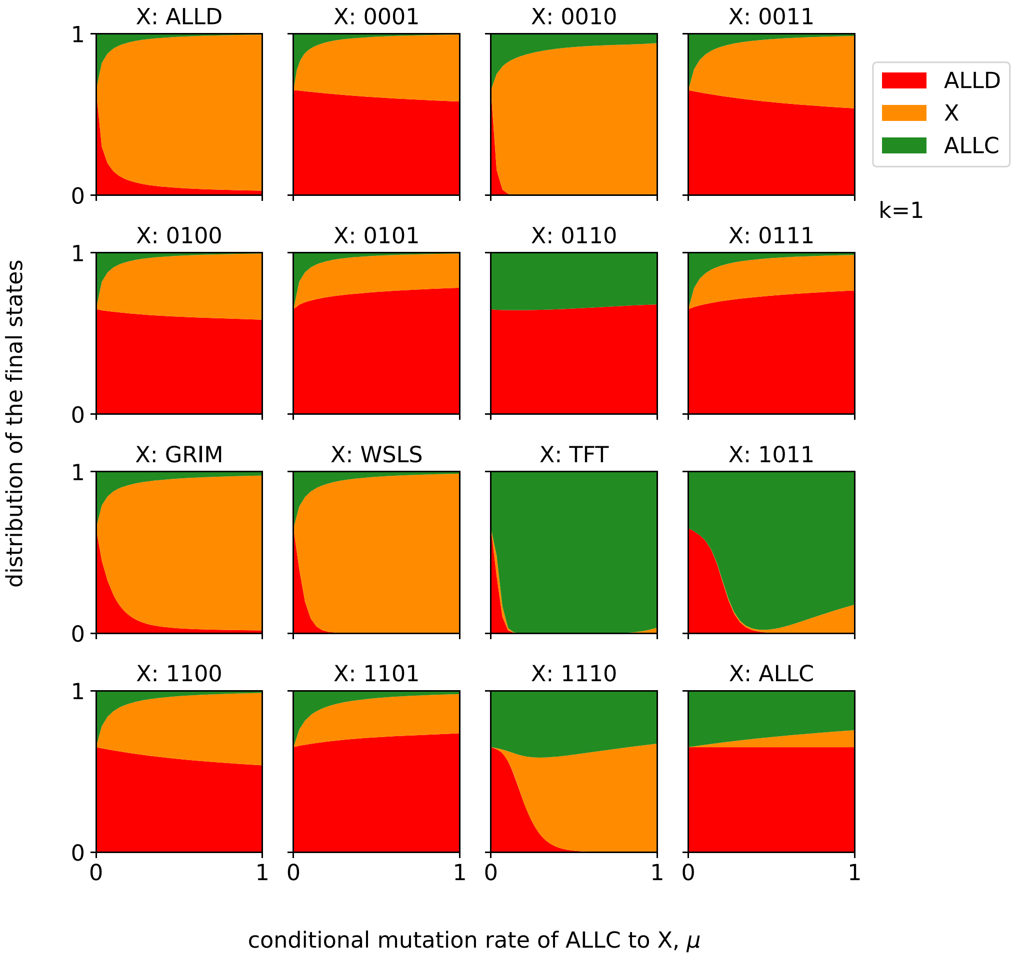


Figure 2. Among all deterministic memory-1 strategies, mutations to TFT result in maximal stabilization of ALLC. We consider $k = 1$ for the stochastic process described in figure 1B. That is, mutations from ALLC to X begin when there is a single ALLD mutant in the population and stops only when the population reaches a state in which ALLD is absent. For each subplot, we consider a fixed memory-1 strategy as X, the strategy to which ALLC conditionally mutates with probability μ. We plot the distribution of the population's three possible final states—where everyone adopts ALLC (green), X (orange) or ALLD (red)—as we vary μ. We present this for all 16 possible deterministic memory-1 strategies as X. Here, a deterministic memory-1 strategy is represented by a four-digit string of 0s and 1s in the form: $p_{CC}p_{CD}p_{DC}p_{DD}$. These elements denote the probability to cooperate in a round when the outcome of the previous round is CC, CD, DC or DD, respectively. The likelihood of restoring the ground state is the highest when ALLC conditionally mutates to TFT (row 3, column 3) at a rate which is optimally high. For this figure, we take a population with 100 individuals. The benefit, *b* and cost of cooperation, *c* in the donation game are 5 and 1, respectively.

We also study the case where rescue response is delayed ($k = 25$). In this case, mutations only occur when the number of ALLD individuals exceed the threshold, $k = 25$. We summarize the result in electronic supplementary material, figure S2. While rescue outcomes are always worse compared to $k = 1$, TFT still acts as the best rescue strategy, among all deterministic memory-1 strategies. Importantly, strategies that performed reasonably well in eradicating ALLD earlier, now do poorly. In fact, for most X, outcomes are almost identical to the case where mutations remain absent ($\mu = 0$). For our parameters, TFT is the only strategy that brings the restoration probability close to one.

We also examine how deterministic memory-1 strategies perform as rescue strategies under an alternative mutational scheme, where the mutation probability is not a step function of the number of ALLD individuals, but increases linearly. Specifically, we consider a mutation rate defined by $\mu = N_D/(N-1)$. In this case, the probability of mutation from ALLC to X is zero when ALLD is absent, increases linearly with the number of ALLD in the population, and reaches one when $N-1$ out of N individuals are ALLD. The results are presented in electronic supplementary material, figure S8. Again, we find that TFT is the best rescue strategy among deterministic memory-1 strategies (see electronic supplementary material, figure S8A). We also compare the performance of the linear mutation function with the step function (electronic supplementary material, figure S8B). For this comparison, we used a step function of $\mu = 0.5$ at $k = 1$. For most rescue strategies, the linear mutation function performed better. But for TFT, there was no notable difference between the two, and for FBF, the step function performed considerably better.

(b) Rescue by reactive strategies

Reactive strategies respond to the co-player's most recent move [15]. They are a two-dimensional subspace of memory-1 strategies. A reactive strategy $S(p,q)$ cooperates with probability p after co-player cooperates, and cooperates with probability q after co-player defects. That is, for a reactive strategy $S(p,q)$, we have $p_{CC} = p_{DC} = p$ and $p_{CD} = p_{DD} = q$. For example, TFT is a reactive strategy with $(p,q) = (1,0)$. We perform numerical computations to identify the best rescue strategy among the reactive strategies (see figure 3A). As before, we identify that rescue is most successful for intermediate mutation rates, μ. In addition, best rescue is performed by a reactive strategy that has $p \approx 1$, and $q \approx 0$. Those strategies are close to TFT. The values of p and q for the best rescue strategy depend on μ. In particular, for some values of μ (these are $\mu = 0.4, 0.5$ and 1 in figure 3A), the best rescue strategy is exactly TFT.

(c) Rescue by memory-1 strategies

We search for the best rescue strategy in the entire memory-1 space, which is given by the hypercube $[0,1]^4$. The details of our search process are described in **Methods**. The findings are summarized in figure 3B. We observe that the best rescue strategy has the form $(1, p_{CD}, p_{DC}, 0)$. Unlike the best rescue strategy in the reactive space, here the best rescue strategy always fully cooperates with ALLC ($p_{CC} = 1$) and fully defects with ALLD ($p_{DD} = 0$). The numerical values of p_{CD} and p_{DC} depend on μ and the remaining model parameters, which are $b = 5$, $c = 1$ and $N = 50$ (see figure 3B). We also plot the self-payoff of the best memory-1 strategy as function of the mutation rate, μ (see figure 3B). For low values of μ, the self-payoff decreases, but for larger values it increases again.

(d) Search for the ideal rescue strategy

Let us now search for the ideal rescue strategy but independent of any constraints that are imposed by specific strategy spaces. The performance of a rescue strategy, X, depends on five numbers: (i) the payoff that X receives from ALLC, (ii) the payoff that X receives from ALLD, (iii) the payoff that ALLC receives from X, (iv) the payoff that ALLD receives from X and (v) the payoff X receives from itself. It is evident from the search within the space of memory-1 strategies that the optimal rescue strategy, X, fully cooperates with ALLC while fully defecting against ALLD. Hence, the first four numbers must be: (i) $b-c$, (ii) 0, (iii) $b-c$, (iv) 0. The only quantity that needs to be optimized is the payoff X receives from itself, which must lie in the interval $[0, b-c]$.

Our results are shown in figure 4B. If mutations are triggered by the presence of a single ALLD individual, $k = 1$, the optimal self-payoff of X decreases with mutation rate, μ, up to a certain point, after which it begins to increase. There are two opposing effects. A high self-payoff of X makes it more likely that ALLD becomes extinct. But a high self-payoff of X makes it harder for ALLC to reach fixation after ALLD has become extinct. For small μ, the first effect is more

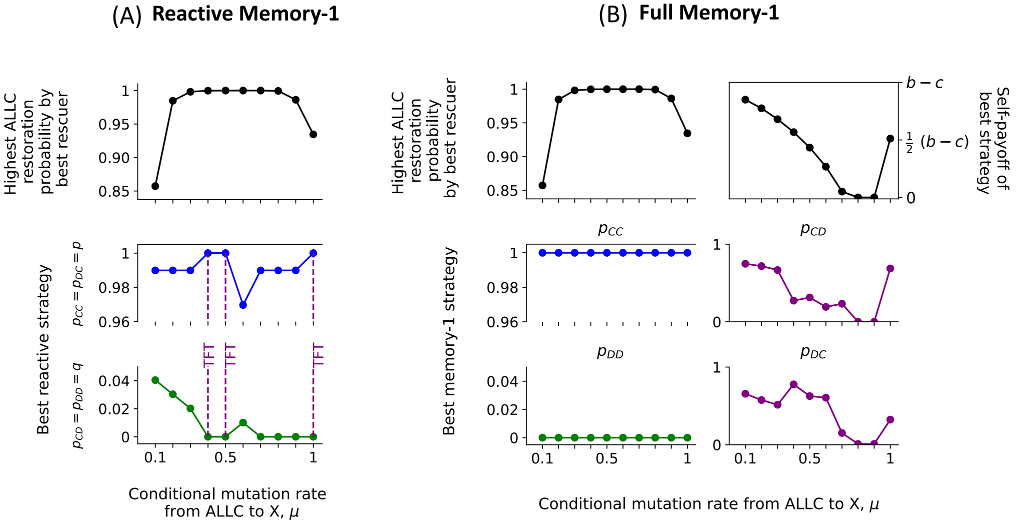


Figure 3. In the reactive memory-1 space, the best rescuers are geometrically close to TFT, whereas in the full memory-1 space, best rescuers have the discriminating property: they fully cooperate with always cooperate and fully defect with always defect. We perform numerical computations to identify the best rescue strategy in the space of reactive memory-1 strategies (A) and in the space of all memory-1 strategies (in B). (A) A reactive memory-1 strategy (p, q) responds solely to co-player's most recent action. It cooperates with probability p after co-player cooperates, or with probability q after co-player defects. That is, $p_{\mathrm{CC}} = p_{\mathrm{DC}} = p$ and $p_{\mathrm{CD}} = p_{\mathrm{DD}} = q$. We examine 500×500 reactive memory-1 strategies uniformly spaced in $[0, 1]^2$. For each X, we compute the restoration probability of the ALLC population. As we vary μ, we plot the maximum restoration probability obtained from 500×500 choices for X, along with the corresponding strategy X that yields this maximum. We find that ALLC restoration probability is highest when μ is neither too high nor too low. Furthermore, the best rescue strategy is geometrically close to $(p, q) = (1, 0)$, TFT. In fact, at $\mu = 0.4$, 0.5 and 1, the best strategy is exactly TFT. (B) We perform a search, similar to a, for finding the best rescue strategy in the entire memory-1 space (for more details on the search process, see *Methods*). The best rescue strategy fully cooperates with ALLC ($p_{\mathrm{CC}} = 1$) and fully defects with ALLD ($p_{\mathrm{DD}} = 0$). The strategy components p_{CD} and p_{DC} of the best rescue strategy determines its self-payoff, which is crucial in the rescue dynamic. We consider $k = 1$, i.e. rescue mutations occur whenever ALLD is present, and a population of size $N = 50$. All other parameters remain consistent with those in figure 2.

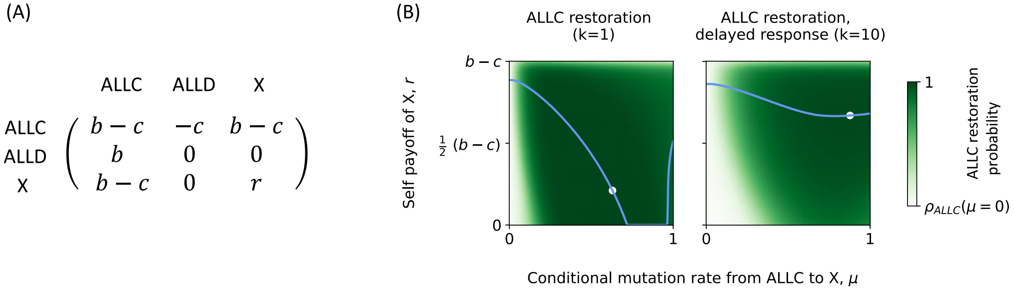
important. For large μ, the second effect is more important. The optimal-r curve aligns with the self-payoff of the best memory-1 rescue strategy of figure 3B, indicating agreement between the two searches.

When rescue is delayed, $k = 10$, the optimal self-payoff, r, is higher. In this case, opposing the spread of ALLD is of primary importance.

5. Evolutionary dynamics in the infinite population model

We now study the evolutionary dynamics of rescue in the limit of infinitely large population size. In the absence of mutation, the standard approach is given by the replicator equation [30–33]. In the presence of mutation, the standard approach is given by the replicator–mutator equation [31,34–37], which we use here. We analyse two versions of the replicator–mutator equation: one in which mutations occur during replication and the other in which mutations occur independently of the replication process [34,38–40]. In the main text, we focus on the latter version. In the electronic supplementary material, we demonstrate that these two approaches produce qualitatively similar outcomes for our context.

In particular, we study the dynamics of a three-strategy system with ALLC, TFT and ALLD. Their relative frequencies are x, y and z, respectively. We have $x + y + z = 1$. Mutations from ALLC



$$\begin{array}{c|ccc} & \text{ALLC} & \text{ALLD} & \text{X} \\ \hline \text{ALLC} & b-c & -c & b-c \\ \text{ALLD} & b & 0 & 0 \\ \text{X} & b-c & 0 & r \end{array}$$

Figure 4. Role of self-payoff of discriminating strategies in the rescue of ALLC. (A) We consider a strategy X that fully cooperates with ALLC, fully defects with ALLD, and receives a payoff of r against itself in an infinitely repeated game, where $r \in [0, b - c]$. In *Methods*, we demonstrate that an example of X is TFT in an environment where two types of implementation errors—specifically, cooperating instead of defecting and defecting instead of cooperating—are rare, independent, and have different probabilities. We are interested to study how r, the self-payoff of X affects the rescue of ALLC. (B) To this end, we numerically compute the probability that the ALLC ground state is restored as we vary r and the conditional mutation probability from ALLC to X, μ. We study two cases, $k = 1$ and $k = 10$. In $k = 1$, mutations occur whenever ALLD is present. In $k = 10$, mutations occur only when the number of ALLD in the population is at least 10. A blue line denotes the values of r that maximize the probability of ground state restoration for each value of μ. The white dot in each subplot denotes the pair (μ, r) that attains the highest restoration probability of ALLC. The lower end of the colour scale denotes the probability of ALLC restoration when mutations do not occur. All other parameters remain consistent with those in figure 3.

to TFT occur at a fixed rate μ. Evolutionary dynamics are given by

$$\left.\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= x(f_x - \bar{f}) - \mu x \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= y(f_y - \bar{f}) + \mu x, \\ \text{and} \qquad \frac{\mathrm{d}z}{\mathrm{d}t} &= z(f_z - \bar{f}). \end{aligned}\right\} \tag{D1}$$

Here, f_x, f_y and f_z represent frequency-dependent fitness of the three types. We have

$$\left.\begin{aligned} f_x &= b(x + y) - c, \\ f_y &= (b - c)(x + y/2) \\ \text{and} \qquad f_z &= bx. \end{aligned}\right\} \tag{5.2}$$

The average fitness is $\bar{f} = xf_x + yf_y + zf_z = (b - c)(x + xy + y^2/2)$. The equation is defined on the simplex Δ_2, which is the set $\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1, 0 \leq x, y, z \leq 1\}$. The simplex Δ_2 is invariant under the dynamics given by equation (D 1). Any point in this set satisfies $\dot{x} + \dot{y} + \dot{z} = 0$. Solutions that begin in this set, remain in this set forever. The corners of the simplex, $(1, 0, 0), (0, 1, 0)$ and $(0, 0, 1)$, represent the homogeneous populations of ALLC, TFT and ALLD, respectively. In proposition 1 of electronic supplementary material, we characterize all rest points of (D 1) that lie in the simplex. In proposition 2, we characterize the conditions under which each rest point is asymptotically stable.

As shown in figure 5A, the evolutionary dynamics depend on the value of the mutation rate, μ. For $\mu = 0$, the corners of the simplex are saddle points. In addition, there is an interior centre, which is surrounded by periodic orbits [41]. In the absence of TFT, ALLC is dominated by ALLD: therefore any mixed population of ALLC and ALLD converges to a homogeneous ALLD population.

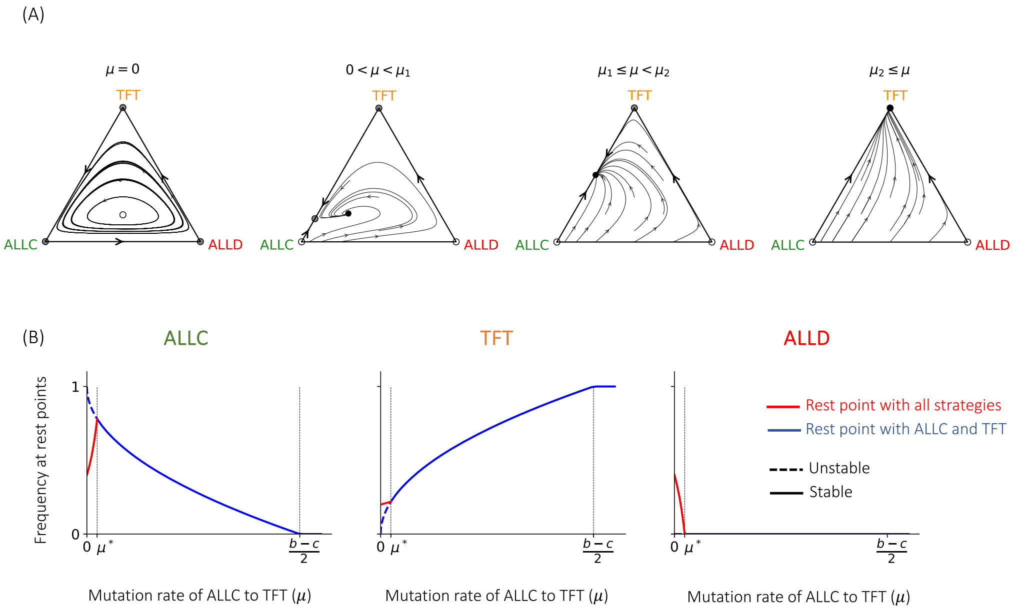


Figure 5. Analysis of the infinite population model with TFT as a rescuer shows an optimal μ ($= \mu_1$) at which ALLC is stable in maximal frequency. (A) The infinite population model demonstrates four unique phase portraits depending on the value of μ. At $\mu = 0$, the dynamic demonstrates a cyclic dominance between ALLC, TFT and ALLD. There is an interior fixed point that acts as centre for periodic orbits. When $0 < \mu < \mu_1$, the dynamic produces a stable coexistence between all three strategies and an unstable coexistence between ALLC and TFT (see equation (5.3) for an expression of μ_1 in terms of b and c). Next, when $\mu \geq \mu_1$ but $\mu < \mu_2 =: (b-c)/2$, the dynamic has a single stable rest point, and in this rest point only TFT and ALLC are present. Finally, for $\mu \geq \mu_2$, the only stable rest point is the homogeneous TFT population. In the last two cases, the stable rest point is the final dynamical fate of all populations that are composed of ALLC and ALLD. (B) We show the frequency of ALLC, TFT and ALLD at rest points, which are either in the interior of the simplex or at the ALLC–TFT edge of the simplex. We mark the frequencies with dashed lines if it corresponds to an unstable rest point and a solid line if it corresponds to a stable rest point. The maximum frequency of ALLC at a stable rest point is at $\mu = \mu_1$. (C) We show how the optimal mutation rate μ_1 varies with cost of cooperation c while we keep the benefit of cooperation b fixed. For this figure, we take the benefit, b and cost of cooperation, c in the donation game as 5 and 1, respectively (for comparison, we make a similar figure in electronic supplementary material, figure S4 for $b = 5$, $c = 3$).

We find two critical mutation rates, μ_1 and μ_2. They are given by

$$\mu_1 = b\left(1 - \sqrt{1 + \frac{c^2}{(b-c)^2}}\right) + \frac{c^2}{b-c} \tag{5.3}$$

and

$$\mu_2 = \frac{b-c}{2}. \tag{5.4}$$

We show in electronic supplementary material that $b > c > 0$ implies that $\mu_1 < \mu_2$.

If $0 < \mu < \mu_1$, there is an unsaturated fixed point on the edge between ALLC and TFT. This saddle point, which attracts all initial conditions where ALLD is absent, can be invaded by ALLD. In addition, there is an interior equilibrium which is asymptotically stable. All initial conditions with $x > 0$ and $z > 0$ converge to the interior equilibrium. All initial conditions with $x = 0$ and $z > 0$ converge to the TFT corner.

If $\mu_1 < \mu < \mu_2$, the fixed point between ALLC and TFT is saturated and asymptotically stable. All initial conditions with $x > 0$ converge to this boundary equilibrium. There is no interior equilibrium. All initial conditions with $x = 0$ and $z > 0$ converge to the TFT corner. For all initial conditions with $z < 1$, ALLD becomes extinct.

If $\mu_2 < \mu$, the corner point of TFT is asymptotically stable. All initial conditions with $x > 0$ or $y > 0$ converge to a population that is homogeneous in TFT. For all initial conditions with $z < 1$, ALLD becomes extinct.

In figure 5B, we show the equilibrium frequencies of the three strategies as functions of the mutation rate μ. We observe that the equilibrium frequency of ALLC attains a maximum for $\mu = \mu_1$. The maximum equilibrium abundance of ALLC for $\mu > 0$ is given by

$$x_{\max} = 1 - \frac{\sqrt{2(b^2 + c^2) - 2b(c + \sqrt{(b-c)^2 + c^2})}}{b - c}. \tag{5.5}$$

Therefore, the deterministic system has an optimum mutation rate that maximizes the abundance of ALLC. In the stochastic system, we also found that intermediate mutation rates were optimal. We also note that $\mu \geq \mu_1$ guarantees the extinction of ALLD.

6. Discussion

In this paper, we have shown that a population of ALLC can be protected against invasion of defectors by mutating toward a rescue strategy. We imagine a situation where the presence of defectors triggers a warning signal that activates the mutation. We examine the resulting mutation-selection dynamics for both a stochastic system, which describes finite population size, and a deterministic system, which describes infinite population size. For both systems, we assume that individuals in the population are randomly paired to play infinitely repeated donation games. Our work adds to the larger body of literature, which studies the role of mutation-generated diversity in the context of evolution of cooperation [42–48].

We have focused on the donation game as it offers the simplest framework for illustrating the role of rescuing strategies. The donation game is widely used for the study of direct reciprocity [16,25]. A donation game is a special case of Prisoner's Dilemma [2], which is the most stringent game for studying evolution of cooperation [3,9]: while mutual cooperation yields higher payoffs for both players than mutual defection, cooperation is not the individually rational choice. Our framework is not limited to the donation game and can readily be applied to other scenarios. In electronic supplementary material, figures S6 and S7, we present the performance of the 16 deterministic memory-1 strategies as rescuers of an ALLC population, when the underlying game is Prisoner's Dilemma or a Stag Hunt. In the Prisoner's Dilemma (see electronic supplementary material, figure S6), TFT remains the best rescue strategy among deterministic memory-1 strategies. In the Stag-Hunt game (electronic supplementary material, figure S7), no deterministic memory-1 strategy reliably restores the ALLC ground state, as the rescue strategy often takes over the population. For example, when GRIM is used, ALLD is suppressed at high mutation rates, but the population is more likely to absorb into the GRIM state than the ALLC state. In such games, employing back-mutations, as previously discussed, may enable the recovery of the ALLC population.

For the stochastic system, we have searched for the optimal rescue strategy in the space of memory-1 strategies. Among deterministic and reactive memory-1 strategies, TFT and TFT-like strategies serve as the best rescuers (figures 2, 3A). In the full, stochastic memory-1 space, the best rescue strategy shares key properties with TFT: it fully cooperates with ALLC but fully defects with ALLD (see figure 3B). For a comparison between TFT and the optimal rescue strategy see electronic supplementary material, figure S10. For such a strategy, its self-payoff determines its efficiency to rescue the ALLC population (see figure 4).

For the deterministic system, we study a replicator–mutator equation, considering the strategies ALLC, ALLD and TFT. We examine how the rate of mutation from ALLC to TFT, μ, affects the three-strategy dynamic (see figure 5). Our analysis reveals a critical mutation rate μ_1 above which ALLD is guaranteed to perish from any mixed population. Furthermore, μ_1 is also optimal: at $\mu = \mu_1$, ALLC attains maximal abundance at an asymptotically stable equilibrium. While we have thoroughly explored TFT as the rescue strategy in the infinite population model, future work will focus on extending the analysis to a more general class of rescue strategies characterized by the following properties: the strategy fully cooperates with ALLC, fully defects against ALLD, and achieves an arbitrary payoff $r \in [0, b - c]$ when playing against itself in direct competition.

Most approaches to evolution of cooperation by direct or indirect reciprocity conclude that cooperation can only be sustained by populations of conditional cooperators (discriminators) [17,24,35,49–54]. Our paper suggests an alternative role for discriminators in evolution. In the context of direct reciprocity, we demonstrate their success as temporary helpers: they only emerge when unconditional cooperators (ALLC) need help to fight off an invasion of defectors. After the defectors are eliminated, the discriminators naturally become extinct thereby restoring the ALLC population. This mechanism of transient helpers is reminiscent of findings from other models of evolution of cooperation—those not based on direct reciprocity—where type X assists in eliminating type Z to protect type Y, but is ultimately eliminated by Y once its purpose is served [55,56].

For our mechanism to succeed, it is important that ALLC dominates the rescue strategy, as this increases the likelihood that, once defectors are eliminated, ALLC can outcompete the helpers and reclaim the population. In our case, TFT, which is a highly effective rescuer, is dominated by ALLC because of its inability to fully cooperate against itself in the presence of implementation errors. In this sense, errors are not an impediment, as previously argued [35,49,51,57], but rather a critical factor behind TFT's role in helping to re-establish stable cooperation. In future works, it will be valuable to explore how the rescue mechanism operates in the context of indirect reciprocity [53,58–63] and in spatial games or evolutionary graph theory [64–67].

7. Methods

(a) Computation of transition matrix

For the finite population model, the transition matrix $\mathbf{T}$ collects probabilities of transition between states in S. The probability of transition from state $\mathbf{N}_a := (i, j, N - i - j)$ to $\mathbf{N}_b := (q, r, N - q - r)$ in a step is given by equation (7.1). In the following, we denote with μ_e the value that the mutation function $\mu_e(N_D; k)$ attains at state $\mathbf{N}_a$ as per equation (3.1). We express the transition probability in terms of the fitnesses $\mathrm{F}_C, \mathrm{F}_D$ and F_X of the three strategies at population state $\mathbf{N}_a$. We define $\mathrm{F}_w := i\mathrm{F}_C + j\mathrm{F}_D + (N - i - j)\mathrm{F}_X$.

$$\mathrm{T}_{\mathbf{N}_a \to \mathbf{N}_b} = \begin{cases} (1-\mu_e) \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{N-i-j}{N} & \text{if } (q,r) = (i+1, j) \\ (1-\mu_e) \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{j}{N} & \text{if } (q,r) = (i+1, j-1) \\ \dfrac{j\mathrm{F}_D}{\mathrm{F}_w} \cdot \dfrac{N-i-j}{N} & \text{if } (q,r) = (i, j+1) \\ \dfrac{(N-i-j)\mathrm{F}_X}{\mathrm{F}_w} \cdot \dfrac{j}{N} + \mu_e \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{j}{N} & \text{if } (q,r) = (i, j-1) \\ \dfrac{j\mathrm{F}_D}{\mathrm{F}_w} \cdot \dfrac{i}{N} & \text{if } (q,r) = (i-1, j+1) \\ \mu_e \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{i}{N} + \dfrac{(N-i-j)\mathrm{F}_X}{\mathrm{F}_w} \cdot \dfrac{i}{N} & \text{if } (q,r) = (i-1, j) \\ 1 - \sum\limits_{\mathbf{s} \neq \mathbf{N}_a} \mathrm{T}_{\mathbf{N}_a \to \mathbf{s}} & \text{if } (q,r) = (i, j) \\ 0 & \text{otherwise} \end{cases}. \tag{7.1}$$

(b) Payoffs when TFT faces ALLC, ALLD or itself in the limit of rare implementation error

Consider the case where strategies are subject to implementation errors. An intended cooperation is executed as defection with probability $\alpha\varepsilon$ and an intended defection is executed as cooperation with probability $\beta\varepsilon$. Let us assume that these two types of error are independent. In this setting, TFT or $(1,0,1,0)$ is effectively $\mathrm{TFT}(\alpha,\beta,\varepsilon) := (1-\alpha\varepsilon, \beta\varepsilon, 1-\alpha\varepsilon, \beta\varepsilon)$. Similarly, ALLD and ALLC are, respectively, $\mathrm{ALLD}(\alpha,\beta,\varepsilon) := (\beta\varepsilon, \beta\varepsilon, \beta\varepsilon, \beta\varepsilon)$, and $\mathrm{ALLC}(\alpha,\beta,\varepsilon) := (1-\alpha\varepsilon, 1-\alpha\varepsilon, 1-\alpha\varepsilon, 1-\alpha\varepsilon)$.

We consider $\alpha, \beta \geq 0$ and compute the expected long-run payoffs for strategy $s_1(\alpha,\beta,\varepsilon)$ against strategy $s_2(\alpha,\beta,\varepsilon)$ in the limit $\varepsilon \to 0^+$. Here, s_1 and s_2 are TFT, ALLD or ALLC. Since these strategies are effectively purely stochastic, the transition matrix $\mathrm{W}(\alpha,\beta,\varepsilon)$, which collects the probability of transition between the states CC, CD, DC and DD between consecutive rounds is irreducible. The Markov chain therefore has a unique stationary distribution, which is given by [68]

$$v(\alpha,\beta,\varepsilon) = (1,1,1,1)\cdot(\mathbb{I} + \mathrm{U} - \mathrm{W}(\alpha,\beta,\varepsilon))^{-1}, \tag{7.2}$$

where $\mathbb{I}$ is the identity matrix of size 4 and U is a 4×4 matrix with all entries equal to 1. The expected long-run payoff of s_1 and s_2 in the limit of rare implementation errors is given by

$$(\pi_1(\alpha,\beta), \pi_2(\alpha,\beta)) = \left(\lim_{\varepsilon\to 0^+} v(\alpha,\beta,\varepsilon)\right)\cdot\begin{pmatrix} b-c & -c & b & 0 \\ b-c & b & -c & 0 \end{pmatrix}^{\mathsf{T}}. \tag{7.3}$$

We compute the payoffs between ALLC, ALLD and TFT using this method. They are

$$\begin{array}{cc} & \begin{array}{ccc} \mathrm{ALLC} & \mathrm{ALLD} & \mathrm{TFT} \end{array} \\ \begin{array}{c} \mathrm{ALLC} \\ \mathrm{ALLD} \\ \mathrm{TFT} \end{array} & \begin{pmatrix} b-c & -c & b-c \\ b & 0 & 0 \\ b-c & 0 & (b-c)\dfrac{\beta}{\alpha+\beta} \end{pmatrix}. \end{array} \tag{7.4}$$

For the case $\alpha = \beta$ (i.e. table 1), TFT earns $(b-c)/2$ against itself. So long as $\alpha + \beta > 0$, TFT can, in principle, earn any self-payoff in the interval $[0, b-c]$. In every calculation in this paper, we have assumed $\alpha = \beta$.

(c) Search process for the best rescue strategy in reactive space

For figure 3A, we perform numerical computations to identify the best rescue strategy, (p, q), in the reactive strategy space, $[0,1]^2$. Specifically, we conduct a grid search over 500×500 uniformly spaced points in $[0,1]^2$ and report the strategy that maximizes the restoration probability of the ground state. For comparison, we perform an additional search restricted to rescue strategies with $p = 1$. We report the comparison in electronic supplementary material, figure S9.

(d) Search process for the best rescue strategy in memory-1 space

For figure 3B, we perform numerical computations to identify the best rescue strategy in the memory-1 space, $[0,1]^4$. For our computations, we consider the following parameters for our model N (population size) $= 50$, b, c (benefit and cost of cooperation) $= 5$, 1 and k (threshold of conditional mutations) $= 1$. Our search process was three-step. In the first step, we performed a $20 \times 20 \times 20 \times 20$ uniformly spaced grid-search in the memory-1 space, and looked for the strategy that yielded highest restoration probability to ALLC. The best strategy from this first step had $p_{\mathrm{CC}} = 1, p_{\mathrm{DD}} = 0$. In the second step, we performed a search in a subset of $[0,1]^4$. In this search, p_{CC} was restricted to $[0.95, 1]$, p_{DD} to $[0, 0.05]$ while p_{CD} and p_{DC} to $[0,1]^2$. We use 20 uniformly spaced points for p_{CC} and p_{DD} and 50 for p_{CD} and p_{DC}. The best strategy from the second search also had $p_{\mathrm{CC}} = 1$ and $p_{\mathrm{DD}} = 0$. In the third and final search, we fix $p_{\mathrm{CC}} = 1, p_{\mathrm{DD}} = 0$ and perform a 500×500 uniformly spaced search in $[0,1]^2$ to find the optimal p_{CD} and p_{DC}.

Data accessibility. The necessary data and code for this manuscript have been archived. They can be found in the referenced link [69]. Data and code are available at [69].

The data are provided in electronic supplementary material [70].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. S.P.: conceptualization, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, writing—review and editing; M.L.: formal analysis, investigation, methodology, software, validation, visualization; M.N.: conceptualization, formal analysis, investigation, methodology, project administration, resources, supervision, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. No funding has been received for this article.

Appendix A. Analysis of the infinite population model

In propositions A.1 and A.2 below, we respectively characterize the fixed points of (D 1) and the corresponding condition for the local stability of these fixed points. Proposition A.4 describes the asymptotic extinction of ALLD from any mixed population.

Upon adding up the dynamical equations in equation (D 1), we get $\dot{x}+\dot{y}+\dot{z}=0$. The triangular simplex $\Delta_2 := \{(x,y,z)\in\mathbb{R}^3 \mid x+y+z=1, x,y,z\geq 0\}$ is thus invariant to (D 1). If the dynamics begins in Δ_2 it remains there forever. We focus on (D 1) in this set. In addition, subsets $\{(x,y,z)\in\Delta_2 \mid z=0\}$ and $\{(x,y,z)\in\Delta_2 \mid x=0\}$ of Δ_2 are also invariant to equation (D 1). These subsets form two edges of the triangular simplex in which ALLD and ALLC are, respectively, absent. Finally, the last edge, $\{(x,0,z) \mid x+z=1\}$ is only invariant if $\mu=0$, otherwise not. The proposition below characterizes all rest points of (D 1) in Δ_2.

Proposition A.1 (Rest points of (D 1) in Δ_2). *For the replicator–mutator equation (D 1), the following are the rest points in Δ_2 provided associated conditions (if any) are met.*

1. $\mathbf{x}_{\mathrm{TFT}} := (0,1,0)$,
2. $\mathbf{x}_{\mathrm{ALLD}} := (0,0,1)$,
3. $\mathbf{x}_{\mathrm{ALLC}} := (1,0,0)$ *if and only if* $\mu=0$,
4. $\mathbf{x}_{\mathrm{E}} := \left(1-\sqrt{\frac{2\mu}{b-c}}, \sqrt{\frac{2\mu}{b-c}}, 0\right)$, *if and only if* $0<\mu<\frac{b-c}{2}$,
5. $\mathbf{x}_{\mathrm{I}} := \left(\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))}, \frac{c+\mu}{b}, 1-\frac{c+\mu}{b}-\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))}\right)$ *iff* $0\leq\mu<\mu_1$.

where μ_1 is given by equation (5.3).

We include proofs of propositions in the section proofs in electronic supplementary material. The above proposition lists five possible rest points of (D 1) in Δ_2. The first three, $\mathbf{x}_{\mathrm{TFT}}$, $\mathbf{x}_{\mathrm{ALLD}}$ and $\mathbf{x}_{\mathrm{ALLC}}$ are three corners of the simplex. In these points, the population only consists of TFT, ALLD or ALLC, respectively. The fourth rest point, $\mathbf{x}_{\mathrm{E}}$, when it exists, is strictly in the interior of the edge where ALLD is absent (i.e. both ALLC and TFT are present). Finally, the last rest point, $\mathbf{x}_{\mathrm{I}}$ when it exists, is strictly in the interior of Δ_2 (i.e. all strategies coexist). We note the following degenerate cases: (a) when $\mu=0$, $\mathbf{x}_{\mathrm{E}}=\mathbf{x}_{\mathrm{ALLC}}$, (b) when $\mu=(b-c)/2$, $\mathbf{x}_{\mathrm{E}}=\mathbf{x}_{\mathrm{TFT}}$. We also note that $\mu_1<\mu_2:=(b-c)/2$ under the assumption that $b>c>0$ (see end of proof of proposition 1). In the proposition below, we characterize the dynamic stability of these rest points.

Proposition A.2 (Asymptotic stability of rest points of (D 1)). *The following statements describe the asymptotic stability of rest points of (D 1) with respect to perturbations in Δ_2. Each statement subsumes that the relevant condition for the fixed point's existence is met (from proposition A.1).*

1. *The rest points, $\mathbf{x}_{\mathrm{ALLD}}$ and $\mathbf{x}_{\mathrm{ALLC}}$, lying at corners of Δ_2, are not stable.*
2. *The third corner rest point, $\mathbf{x}_{\mathrm{TFT}}$, is stable if and only if $\mu>(b-c)/2$.*

3. *The rest point* $\mathbf{x}_{\mathrm{E}}$, *which lies on the boundary where ALLD is absent, is stable if and only if* $\mu > \mu_1$.
4. *The rest point*, $\mathbf{x}_{\mathrm{I}}$, *lying strictly in the interior of* Δ_2, *is stable if and only if* $\mu > 0$.

Furthermore, if the fixed point $\mathbf{x}_{\mathrm{E}}$ *exists, any solution starting from the set* $\{(x, 1-x, 0) \mid 0 < x \leq 1\}$ *asymptotically converges to* $\mathbf{x}_{\mathrm{E}}$ *(the set includes* $x = 1$ *if* $\mu > 0$*, otherwise not). In addition, at* $\mu = 0$*, the interior fixed point* $\mathbf{x}_{\mathrm{I}}$ *acts as a centre for periodic orbits in* Δ_2*.*

First, the proposition claims that rest points corresponding to homogeneous populations of ALLC and ALLD are not stable. We show that introducing TFT in small amounts to an ALLD population destabilizes it. Similarly, introducing ALLD destabilizes an ALLC population. When $\mu \leq (b-c)/2$, a homogeneous TFT population is also not stable since a small introduction of ALLC grows. However, if mutations are too frequent, $\mu > (b-c)/2$, a homogeneous TFT population is stable. Second, the proposition claims that the value of μ determines the stability of the ALLC–TFT coexistence. This coexistence is stable when $\mu_1 < \mu < (b-c)/2$. When $\mu \leq \mu_1$, this coexistence is not stable; a small introduction of ALLD into this mixed population grows. Finally, the proposition claims that the three types—ALLC, TFT and ALLD—can stably coexist, if conditions allow for such a coexistence in the first place *and* the mutation rate is positive (that is, $0 < \mu < \mu_1$). At $\mu = 0$, this coexistence acts as a centre for periodic orbits. We also show that any population containing ALLC but no ALLD asymptotically reaches this ALLC–TFT coexistence. This also implies that this coexistence is always stable to perturbations that do not involve ALLD.

Corollary A.3. *The strategy ALLC attains maximal frequency at a stable coexistence when* $\mu = \mu_1$.

Proposition A.4 (Extinction of ALLD from any mixed population when μ is sufficiently high). *If* $\mu \geq \mu_1$, *any solution* $(x(t), y(t), z(t))$ *of* (D 1) *with* $(x(0), y(0), z(0)) \in \Delta_2$ *such that* $z(0) \in [0, 1)$ *has* $\lim_{t\to\infty} z(t) = 0$.

This proposition claims that the dynamics (D 1) eliminates ALLD from any starting population that has ALLD and other strategies (either TFT, ALLC or both), provided the mutation rate, μ is at least μ_1.

References

1. Trivers RL. 1971 The evolution of reciprocal altruism. *Q Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)
2. Axelrod R. 1984 *The evolution of cooperation*. New York, NY: Basic Books.
3. Nowak MA. 2006 Five rules for the evolution of cooperation. *Science* **314**, 1560–1563. (doi:10.1126/science.1133755)
4. Wilson DS, Wilson EO. 2007 Rethinking the theoretical foundation of sociobiology. *Q Rev. Biol.* **82**, 327–348. (doi:10.1086/522809)
5. Dugatkin LA. 2002 Cooperation in animals: an evolutionary overview. *Biol. Phil.* **17**, 459–476. (doi:10.1023/A:1020573415343)
6. Riolo RL, Cohen MD, Axelrod R. 2001 Evolution of cooperation without reciprocity. *Nature* **414**, 441–443. (doi:10.1038/35106555)
7. Zaggl MA. 2014 Eleven mechanisms for the evolution of cooperation. *J. Inst. Econ.* **10**, 197–230.
8. Traulsen A, Nowak MA. 2006 Evolution of cooperation by multilevel selection. *Proc. Natl Acad. Sci. USA* **103**, 10 952–10 955. (doi:10.1073/pnas.0602530103)
9. Nowak MA *et al.* 2012 Evolving cooperation. *J. Theor. Biol.* **299**, 1–8. (doi:10.1016/j.jtbi.2012.01.014)
10. Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. 2017 Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51. (doi:10.1016/j.physrep.2017.05.004)
11. Perc M, Szolnoki A. 2010 Coevolutionary games–a mini review. *BioSystems* **99**, 109–125. (doi:10.1016/j.biosystems.2009.10.003)
12. Fu F, Hauert C, Nowak MA, Wang L. 2008 Reputation-based partner choice promotes cooperation in social networks. *Phys. Rev. E–Stat., Nonlinear, Soft Matter Phys.* **78**, 026117. (doi:10.1103/PhysRevE.78.026117)
13. Milinski M. 1987 Tit for tat in sticklebacks and the evolution of cooperation. *Nature* **325**, 433–435. (doi:10.1038/325433a0)

14. Fundenberg D, Maskin E. 1990 Evolution and cooperation in noisy repeated games. *Am. Econ. Rev.* **80**, 274–279.
15. Nowak MA, Sigmund K. 1992 Tit for tat in heterogeneous populations. *Nature* **355**, 250–253. (doi:10.1038/355250a0)
16. Van Veelen M, García J, Rand DG, Nowak MA. 2012 Direct reciprocity in structured populations. *Proc. Natl Acad. Sci. USA* **109**, 9929–9934. (doi:10.1073/pnas.1206694109)
17. Stewart AJ, Plotkin JB. 2013 From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc. Natl Acad. Sci. USA* **110**, 15 348–15 353. (doi:10.1073/pnas.1306246110)
18. Hilbe C, Chatterjee K, Nowak MA. 2018 Partners and rivals in direct reciprocity. *Nat. Hum. Behav.* **2**, 469–477. (doi:10.1038/s41562-018-0320-9)
19. Chen X, Fu F. 2023 Outlearning extortioners: unbending strategies can foster reciprocal fairness and cooperation. *PNAS Nexus* **2**, pgad176. (doi:10.1093/pnasnexus/pgad176)
20. Fudenberg D, Maskin E. 2009 The folk theorem in repeated games with discounting or with incomplete information. In *A Long-run collaboration on long-run games*, pp. 209–230. World Scientific.
21. Hilbe C, Traulsen A, Sigmund K. 2015 Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games Econ. Behav.* **92**, 41–52. (doi:10.1016/j.geb.2015.05.005)
22. Press WH, Dyson FJ. 2012 Iterated Prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl Acad. Sci. USA* **109**, 10 409–10 413. (doi:10.1073/pnas.1206569109)
23. Wedekind C, Milinski M. 1996 Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proc. Natl Acad. Sci. USA* **93**, 2686–2689. (doi:10.1073/pnas.93.7.2686)
24. 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. (doi:10.1038/364056a0)
25. Sigmund K. 2010 *The calculus of selfishness*. Princeton, NJ: Princeton University Press.
26. Nowak M, Sigmund K. 1993 Chaos and the evolution of cooperation. *Proc. Natl Acad. Sci. USA* **90**, 5091–5094. (doi:10.1073/pnas.90.11.5091)
27. Nowak MA. 2006 *Evolutionary dynamics: exploring the equations of life*. Cambridge, MA: Harvard University Press.
28. Wu B, Bauer B, Galla T, Traulsen A. 2015 Fitness-based models and pairwise comparison models of evolutionary games are typically different–even in unstructured populations. *New J. Phys.* **17**, 023043. (doi:10.1088/1367-2630/17/2/023043)
29. Zagorsky BM, Reiter JG, Chatterjee K, Nowak MA. 2013 Forgiver triumphs in alternating prisoner's dilemma. *PLoS ONE* **8**, e80814. (doi:10.1371/journal.pone.0080814)
30. Taylor PD, Jonker LB. 1978 Evolutionary stable strategies and game dynamics. *Math. Biosci.* **40**, 145–156. (doi:10.1016/0025-5564(78)90077-9)
31. Hofbauer J, Sigmund K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge university press.
32. Hofbauer J, Sandholm WH. 2011 Survival of dominated strategies under evolutionary dynamics. *Theor. Econ.* **6**, 341–377. (doi:10.3982/TE771)
33. Sandholm WH. 2010 *Population games and evolutionary dynamics*. Cambridge, MA: MIT press.
34. Nowak MA, Komarova NL, Niyogi P. 2001 Evolution of universal grammar. *Science* **291**, 114–118. (doi:10.1126/science.291.5501.114)
35. Imhof LA, Fudenberg D, Nowak MA. 2005 Evolutionary cycles of cooperation and defection. *Proc. Natl Acad. Sci. USA* **102**, 10 797–10 800. (doi:10.1073/pnas.0502589102)
36. Komarova NL. 2004 Replicator–mutator equation, universality property and population dynamics of learning. *J. Theor. Biol.* **230**, 227–239. (doi:10.1016/j.jtbi.2004.05.004)
37. Duong MH, Han TA. 2020 On equilibrium properties of the replicator–mutator equation in deterministic and random games. *Dyn. Games Appl.* **10**, 641–663. (doi:10.1007/s13235-019-00338-8)
38. FP. Toupo D, Strogatz SH. 2015 Nonlinear dynamics of the rock-paper-scissors game with mutations. *Phys. Rev. E* **91**, 052907. (doi:10.1103/PhysRevE.91.052907)
39. Mobilia M. 2010 Oscillatory dynamics in rock–paper–scissors games with mutations. *J. Theor. Biol.* **264**, 1–10. (doi:10.1016/j.jtbi.2010.01.008)
40. Toupo DFP, Rand DG, Strogatz SH. 2014 Limit cycles sparked by mutation in the repeated Prisoner's dilemma. *Int. J. Bifurcation Chaos* **24**, 1430035. (doi:10.1142/S0218127414300353)
41. Nowak M, Sigmund K. 1989 Game-dynamical aspects of the prisoner's dilemma. *Appl. Math. Comput.* **30**, 191–213. (doi:10.1016/0096-3003(89)90052-0)

42. Kandori M, Mailath GJ, Rob R. 1993 Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56. (doi:10.2307/2951777)
43. Eriksson A, Lindgren K. 2005 Cooperation driven by mutations in multi-person prisoner's dilemma. *J. Theor. Biol.* **232**, 399–409. (doi:10.1016/j.jtbi.2004.08.020)
44. Helbing D, Szolnoki A, Perc M, Szabó G. 2010 Defector-accelerated cooperativeness and punishment in public goods games with mutations. *Phys. Rev. E* **81**, 057104. (doi:10.1103/PhysRevE.81.057104)
45. Spichtig M, Egas M. 2019 When and how does mutation-generated variation promote the evolution of cooperation? *Games* **10**, 4. (doi:10.3390/g10010004)
46. Ichinose G, Satotani Y, Sayama H. 2018 How mutation alters the evolutionary dynamics of cooperation on networks. *New J. Phys.* **20**, 053049. (doi:10.1088/1367-2630/aac2a7)
47. Feng K, Han S, Feng M, Szolnoki A. 2024 An evolutionary game with reputation-based imitation-mutation dynamics. *Appl. Math. Comput.* **472**, 128618. (doi:10.1016/j.amc.2024.128618)
48. Tkadlec J, Hilbe C, Nowak MA. 2023 Mutation enhances cooperation in direct reciprocity. *Proc. Natl Acad. Sci. USA* **120**, e2221080120. (doi:10.1073/pnas.2221080120)
49. Rand DG, Ohtsuki H, Nowak MA. 2009 Direct reciprocity with costly punishment: generous tit-for-tat prevails. *J. Theor. Biol.* **256**, 45–57. (doi:10.1016/j.jtbi.2008.09.015)
50. 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)
51. Brandt H, Sigmund K. 2006 The good, the bad and the discriminator–errors in direct and indirect reciprocity. *J. Theor. Biol.* **239**, 183–194. (doi:10.1016/j.jtbi.2005.08.045)
52. Glynatsi NE, Nowak MA, Hilbe C. 2024 Conditional cooperation with longer memory. arXiv preprint (http://arxiv.org/abs/2402.02437)
53. Nowak MA, Sigmund K. 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298. (doi:10.1038/nature04131)
54. Nowak MA, Sigmund K. 1998 Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577. (doi:10.1038/31225)
55. Szolnoki A, Szabó G, Czakó L. 2011 Competition of individual and institutional punishments in spatial public goods games. *Phys. Rev. E* **84**, 046106. (doi:10.1103/PhysRevE.84.046106)
56. Lee HW, Cleveland C, Szolnoki A. 2022 When costly migration helps to improve cooperation. *Chaos: Interdiscip. J. Nonlinear Sci.* **32**, 093103. (doi:10.1063/5.0100772)
57. Nowak MA, Sasaki A, Taylor D, Fudenberg C. 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650. (doi:10.1038/nature02414)
58. Nowak MA, Sigmund K. 1998 The dynamics of indirect reciprocity. *J. Theor. Biol.* **194**, 561–574. (doi:10.1006/jtbi.1998.0775)
59. Ohtsuki H, Iwasa Y. 2006 The leading eight: social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* **239**, 435–444. (doi:10.1016/j.jtbi.2005.08.008)
60. Santos FP, Santos FC, Pacheco JM. 2018 Social norm complexity and past reputations in the evolution of cooperation. *Nature* **555**, 242–245. (doi:10.1038/nature25763)
61. Hilbe C, Schmid L, Tkadlec J, Chatterjee K, Nowak MA. 2018 Indirect reciprocity with private, noisy, and incomplete information. *Proc. Natl Acad. Sci. USA* **115**, 12 241–12 246. (doi:10.1073/pnas.1810565115)
62. Schmid L, Chatterjee K, Hilbe C, Nowak MA. 2021 A unified framework of direct and indirect reciprocity. *Nat. Hum. Behav.* **5**, 1292–1302. (doi:10.1038/s41562-021-01114-8)
63. Pal S, Hilbe C, Glynatsi NE. 2024 The co-evolution of direct, indirect and generalized reciprocity. arXiv preprint (http://arxiv.org/abs/2411.03488)
64. Nowak MA, May RM. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)
65. Lieberman E, Hauert C, Nowak MA. 2005 Evolutionary dynamics on graphs. *Nature* **433**, 312–316. (doi:10.1038/nature03204)
66. Ohtsuki H, Hauert C, Lieberman E, Nowak MA. 2006 A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505. (doi:10.1038/nature04605)
67. Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, Nowak MA. 2017 Evolutionary dynamics on any population structure. *Nature* **544**, 227–230. (doi:10.1038/nature21723)
68. Couto MC, Pal S. 2023 Introspection dynamics in asymmetric multiplayer games. *Dyn. Games Appl.* **13**, 1256–1285. (doi:10.1007/s13235-023-00525-8)
69. Pal S, Lambert M, Nowak MA. 2025 Stabilizing unconditional cooperation. (doi:10.5281/zenodo.14246418)
70. Pal S, Lambert M, Nowak MA. 2025 Stabilizing unconditional cooperation. Figshare. (doi:10.6084/m9.figshare.c.7879882)