Article

Evolution of cooperation through cumulative reciprocity

Juan Li ⬤[1,2], Xiaowei Zhao[1,3], Bing Li[1], Charlotte S. L. Rossetti[4], Christian Hilbe ⬤[4,5] ✉ and Haoxiang Xia ⬤[1,2,5] ✉

Reciprocity is a simple principle for cooperation that explains many of the patterns of how humans seek and receive help from each other. To capture reciprocity, traditional models often assume that individuals use simple strategies with restricted memory. These memory-1 strategies are mathematically convenient, but they miss important aspects of human reciprocity, where defections can have lasting effects. Here we instead propose a strategy of cumulative reciprocity. Cumulative reciprocators count the imbalance of cooperation across their previous interactions with their opponent. They cooperate as long as this imbalance is sufficiently small. Using analytical and computational methods, we show that this strategy can sustain cooperation in the presence of errors, that it enforces fair outcomes and that it evolves in hostile environments. Using an economic experiment, we confirm that cumulative reciprocity is more predictive of human behaviour than several classical strategies. The basic principle of cumulative reciprocity is versatile and can be extended to a range of social dilemmas.

Evolutionary game theory provides a formal framework to study the evolution of cooperation, which is a far-reaching problem that has attracted great attention[1–3]. The simplest and most widely used model to study this problem is the prisoner's dilemma[2,4]. In the prisoner's dilemma, two individuals independently decide whether to cooperate. Mutual cooperation is optimal for the pair, yet each individual is tempted to defect. Although the basic premise of the game is simple, it approximates the logic of many cooperative interactions in biological, societal and artificial worlds, including friends who exchange favours[5], animals who exchange food or other services[6] and nations that coordinate their policies[7]. When there is only a single round of the prisoner's dilemma, defection is the only Nash equilibrium. However, if individuals interact repeatedly, they can sustain cooperation through conditionally cooperative strategies[8]. The respective mechanism of cooperation is called direct reciprocity[1,3].

The key to sustaining cooperation in a repeated prisoner's dilemma is to act like a 'partner'[9]. As long as the opponent is cooperative, a partner should go along. However, once an opponent defects, a partner strategy needs to make sure that the opponent cannot gain a lasting advantage. Examples of such partner strategies are tit-for-tat (TFT)[2], generous tit-for-tat (GTFT)[10,11] and win-stay lose-shift (WSLS)[12,13], among many others[14–21]. Although each of these strategies can succeed in certain environments[2,10,13], they also have well-known weaknesses. For example, TFT is unable to sustain cooperation in the presence of errors[22]. GTFT typically fails to evolve when individuals have access to a richer strategy space[23]. Finally, WSLS is stable only when the benefit of mutual cooperation is sufficiently large[13]. The problem of identifying successful strategies of direct reciprocity becomes even more complex when interactions take place among more than two individuals[14,24], or when the benefit of cooperation can change in time[25,26].

Traditionally, much of the existing work on direct reciprocity has been confined to individuals with restricted memory. The most common assumption is that all individuals have one-round memory[27], but researchers have also identified several promising strategies that take into account the last two or three rounds[20,21,28]. While some of these strategies exhibit remarkable robustness properties (especially

[1]Institute of Systems Engineering, Dalian University of Technology, Dalian, China. [2]Center for Big Data and Intelligent Decision-Making, Dalian University of Technology, Dalian, China. [3]School of Software Technology, Dalian University of Technology, Dalian, China. [4]Max Planck Institute for Evolutionary Biology, Plön, Germany. [5]These authors contributed equally: Christian Hilbe and Haoxiang Xia. ✉e-mail: hilbe@evolbio.mpg.de; hxxia@dlut.edu.cn
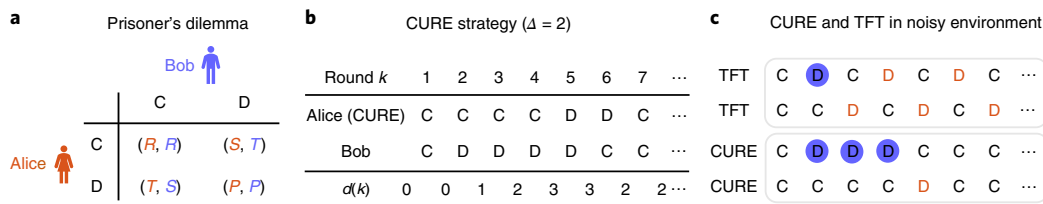
a Prisoner's dilemma

Bob

		C	D
Alice	C	(R, R)	(S, T)
	D	(T, S)	(P, P)

b CURE strategy ($\Delta = 2$)

Round k	1	2	3	4	5	6	7	$\cdots$
Alice (CURE)	C	C	C	C	D	D	C	$\cdots$
Bob	C	D	D	D	D	C	C	$\cdots$
$d(k)$	0	0	1	2	3	3	2	$\cdots$

c CURE and TFT in noisy environment

TFT	C	D	C	D	C	D	C	$\cdots$
TFT	C	C	D	C	D	C	D	$\cdots$
CURE	C	D	D	D	C	C	C	$\cdots$
CURE	C	C	C	C	D	C	C	$\cdots$

Fig. 1 | CURE in the repeated prisoner's dilemma. a, The payoff matrix of a prisoner's dilemma game. For simulations, we use $R = 3$, $S = 0$, $T = 5$ and $P = 1$, unless specified otherwise. **b**, CURE defects when the defection difference statistic $d(k)$ exceeds the tolerance threshold (in this case, $\Delta = 2$). Otherwise, CURE keeps cooperating. **c**, When two TFT players meet and one player makes an error, cooperation breaks down. When two CURE players meet, cooperation is robust even when one player mistakenly defects in three of the rounds ($\Delta = 2$). The blue circle indicates the occurrence of an error, and the red letter indicates the defection caused by the error.

when errors can be assumed to be vanishingly rare), strategies with a pre-defined memory length often fail to capture certain important aspects of human behaviour. For example, individuals might often find it easier to forgive a defecting opponent if this opponent is generally cooperative. To encode these more nuanced responses, individuals need to resort to an opponent's cumulative behaviour, during the entire course of their previous interactions.

Perhaps the most natural way to introduce such behaviours is to let individuals count how often each of them has defected so far. To formalize this idea, consider a repeated prisoner's dilemma between two individuals named 'Alice' and 'Bob'. Suppose that, during their first $k-1$ interactions, Alice defected in n_A rounds whereas Bob defected in n_B rounds. Alice may choose to cooperate in round k unless Bob defected substantially more often in the past than she did. In other words, Alice would cooperate unless $n_B - n_A > \Delta_A$, where $\Delta_A \geq 0$ can be interpreted as Alice's tolerance level. Herein, we refer to this kind of strategy as cumulative reciprocity (CURE).

While CURE is straightforward to define, analytical results are more difficult to obtain, compared with the case of memory-1 strategies. Nevertheless, such results are feasible. First, we show that CURE is indeed a partner strategy in the absence of implementation errors. Second, similar to previous work on zero-determinant strategies[29–34], individuals can use CURE to enforce fairness. If one player is a cumulative reciprocator, both players are guaranteed to get the same payoff, independent of the opponent's strategy. Third, even in the presence of (rare) errors, the payoff of CURE against itself is approximately optimal. At the same time, unconditional defectors cannot invade. Further simulations suggest that individuals are most likely to adopt CURE when most well-known memory-1 strategies fail. We further support these theoretical findings with a behavioural experiment. According to this experiment, CURE is better able to explain human behaviour than many classical strategies such as TFT[2], WSLS[12,13] or previously proposed memory-k strategies[20,21,28].

Overall, our findings suggest that, when cooperation is particularly costly, simple strategies based on an opponent's last behaviour do not suffice. In such environments, it takes a more cumulative assessment of the players' past actions to sustain cooperation. Here, we combine various mathematical, computational and experimental methods to facilitate the analysis of CURE. Although many of our analyses focus on the classical prisoner's dilemma, the main principles of CURE extend to multiplayer interactions, as well as to stochastic games in which payoffs fluctuate in time. In all these applications, CURE proves to be a simple mechanism to sustain fairness and cooperation.

Results
The repeated prisoner's dilemma with CURE
We first consider pairwise interactions between only two individuals (Fig. 1a). We refer to the players as Alice and Bob. In each round, both Alice and Bob independently decide whether they want to cooperate or defect. If they both cooperate (denoted by CC), they get the reward R.

If they defect (DD), they both receive the punishment payoff P. Finally, if one player cooperates while the other defects (CD or DC), the cooperator gets the sucker's payoff S whereas the defector gets the temptation payoff T. Payoffs satisfy the characteristic conditions of the prisoner's dilemma, $S < P < R < T$ and $2R > T + S$. That is, mutual cooperation is the best outcome for the pair, yet each player individually prefers to defect. For our numerical simulations, we use the payoffs of Axelrod[2] ($R = 3$, $S = 0$, $T = 5$ and $P = 1$), unless stated otherwise.

The players' actions may be subject to 'trembling hand' errors[22,35]. That is, a player who intends to cooperate may instead defect with some probability $0 \leq \varepsilon < 1/2$ (similarly, a player who wishes to defect may cooperate with the same probability). For analytical results, we suppose that the game is repeated infinitely. After each round, there is another round. We complement these analytical results with simulations for long but finitely repeated games. The overall payoffs of Alice and Bob are defined as their expected payoffs per round. For details, see Methods.

Traditionally, much of the work on reciprocity assumes that players either have finite recall[20,36–39] or that their decisions can be encoded with finite state automata[40,41]. Instead, here, we propose a strategy of cumulative reciprocity. We formalize this strategy by introducing two counter variables, $n_A(k)$ and $n_B(k)$. These variables record how often Alice and Bob have defected before round k. Let $d(k)$ denote the difference, $d(k) = n_B(k) - n_A(k)$. We refer to $d(k)$ as the defection difference statistic. We say that Alice adopts the strategy CURE if, in any given round k, she cooperates if and only if this defection difference statistic is below a pre-defined threshold $d(k) \leq \Delta_A$. We interpret $\Delta_A \geq 0$ as Alice's tolerance level. If it is zero, Alice demands that Bob be at least as cooperative as Alice. For larger values of Δ_A, Alice becomes increasingly more lenient. Figure 1b depicts the basic logic of CURE.

In contrast to memory-1 strategies, a cumulative reciprocator takes the entire history of the game into account. As a result, the mathematics becomes more intricate. Whereas games between two memory-1 players can be represented as a Markov chain with four possible states (the possible outcomes of any given round), the state variable $d(k)$ of a cumulative reciprocator can assume arbitrary integer values. Perhaps somewhat surprisingly, it is still possible to derive analytical results. To this end, we represent the dynamics between a CURE player and its opponent by an infinitely dimensional linear system. In many important cases, this system can be solved. We summarize our results in the following. All details and proofs are in Supplementary Information Section 1.

Payoffs against selected strategies
To gain some first insights into the performance of CURE, we first study games between two CURE players. In case they both use the same tolerance level $\Delta \geq 1$, each player's average cooperation rate ρ_{CURE} becomes

$$\rho_{CURE} = 1 - \frac{2 - 3\varepsilon + 2\Delta(1 - 2\varepsilon)}{1 - 2\varepsilon^2 + 2\Delta(1 - 2\varepsilon)}\varepsilon. \qquad (1)$$
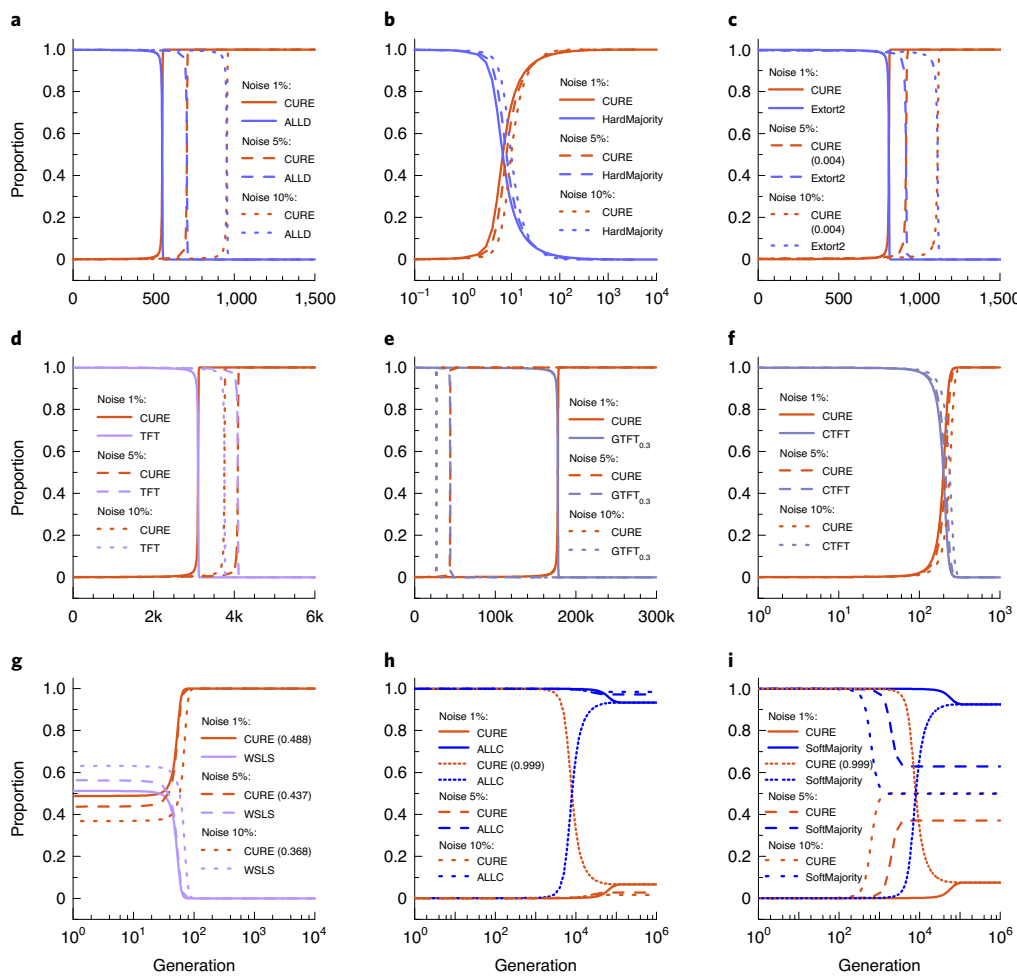
Fig. 2 | The evolutionary dynamics of pairwise strategy competitions. We explore whether CURE can emerge when initially rare. To this end, we explore the evolutionary dynamics for nine different resident populations and three different noise rates (1%, 5% and 10%). Initially, the fraction of cumulative reciprocators is set to 0.001. We explore the further dynamics with simulations. **a**–**f**, Results for CURE versus resident strategies with little cooperation: ALLD (**a**), HardMajority (**b**) and an extortionate strategy (**c**) and versus different variations of tit-for-tat: TFT (**d**), GTFT$_{0.3}$ (**e**) and CTFT (**f**). All these resident populations are invaded by CURE. **g**, Against WSLS, the critical initial frequency of CURE to invade is approximately 0.488 at the noise rate of 1%, 0.437 at the noise rate of 5% and 0.368 at the noise rate of 10%. **h,i**, Finally, for two resident populations which tend to be highly cooperative: ALLC (**h**) and SoftMajority (**i**), CURE does not invade. The general parameters are the same as before. For GTFT, we use a generosity parameter of 0.3 (that is, after a defection, a GTFT player cooperates with 30% probability). For the extortionate strategy, we use an extortion factor of 2 (that is, this extortionate strategy strives to get twice of the payoff surplus of its co-player[29]). For details on the strategies used, see Supplementary Information Section 2.

In particular, the cooperation rate approaches 1 as errors become rare, $\varepsilon \to 0$. Therefore, although CURE follows a similar basic principle to TFT, it is much more robust with respect to noise[27] (Fig. 1c). Overall, the resulting average payoff π (CURE, CURE) is

$$\pi(\text{CURE, CURE}) = \frac{(1-\varepsilon)^2(1-2(1-\varepsilon)\varepsilon+2\Delta(1-2\varepsilon))}{1-2\varepsilon^2+2\Delta(1-2\varepsilon)} \times R$$
$$+ \frac{2\varepsilon(1-\varepsilon)\left((1-\varepsilon)^2+\Delta(1-2\varepsilon)\right)}{1-2\varepsilon^2+2\Delta(1-2\varepsilon)} \times (T+S) \quad (2)$$
$$+ \frac{\varepsilon^2(3-2(3-\varepsilon)\varepsilon+2\Delta(1-2\varepsilon))}{1-2\varepsilon^2+2\Delta(1-2\varepsilon)} \times P.$$

As one may expect, this payoff is monotonically decreasing in the error rate, but increasing in the strategy's tolerance level. The more tolerant the two players are, the better they can cope with each other's unintentional errors. Analogous formulas can also be derived if the two cumulative reciprocators use different thresholds Δ_A and Δ_B.

Of course, a strategy's payoff against itself is only one possible measure of a strategy's ability to sustain cooperation. In a next step, we consider interactions between CURE and other well-known strategies. We start by matching a cumulative reciprocator with an unconditional defector (ALLD). We find that, in that case, both players' cooperation rates assume the theoretical minimum ε (which is also how often ALLD cooperates against itself). The result is intuitive. If Alice adopts CURE and Bob adopts ALLD, Alice is only expected to cooperate in the first Δ_A rounds. From then on, both players effectively implement an ALLD strategy. At that point, either of them only cooperates in case of an error. Because the first Δ_A rounds are negligible in infinitely repeated games, both players obtain the same payoff. As a result, we find that CURE weakly dominates ALLD for all error rates and all tolerance levels (Supplementary Information Section 1). Cumulative reciprocators can therefore cope with unconditional defectors even better than TFT (TFT never dominates ALLD in the presence of errors[22]). Similarly, we can also consider games between cumulative reciprocators and unconditional cooperators (ALLC). Here, cooperation rates are maximal, $1-\varepsilon$, which is
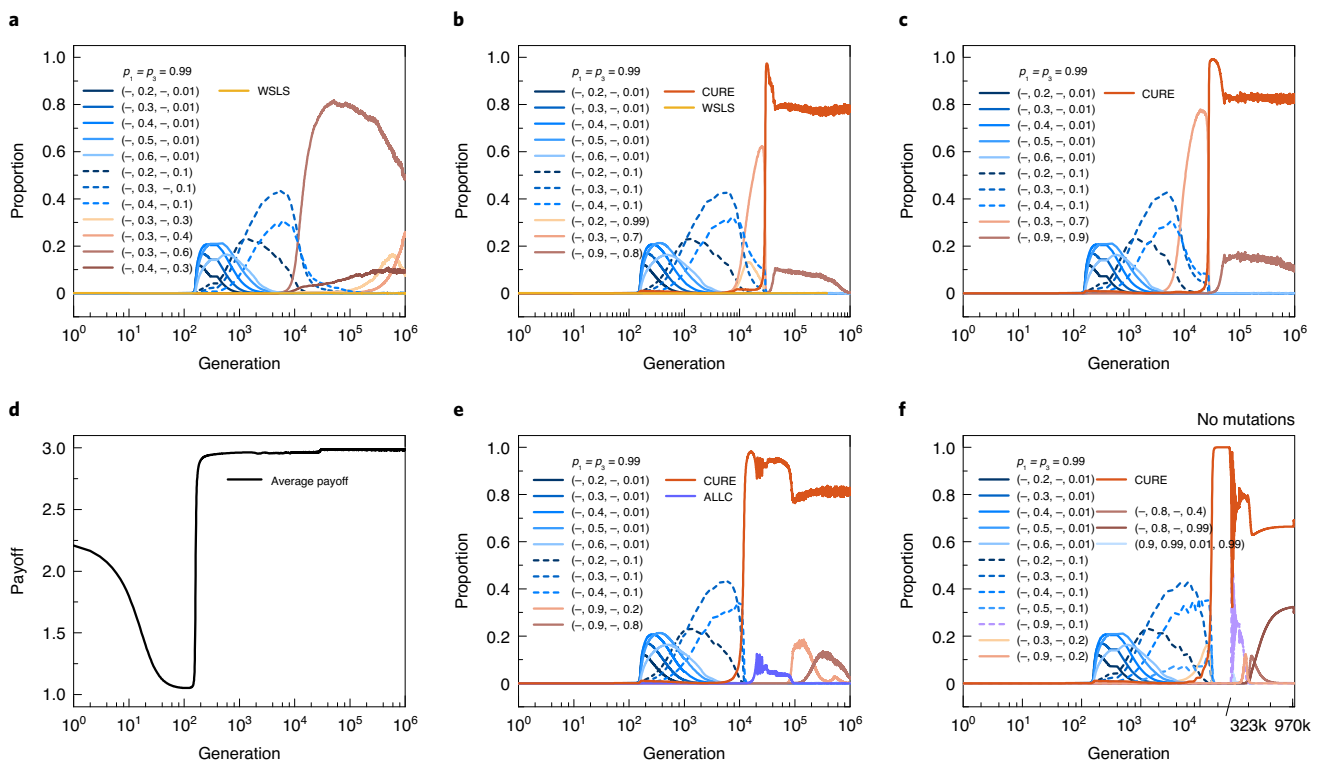
Fig. 3 | Evolution of CURE in populations of memory-1 players. We consider the evolutionary dynamics when individuals can choose among a general set of memory-1 strategies and CURE. Initially, all available strategies are equally abundant in the population. As the simulation proceeds, we first often observe the emergence of TFT-like strategies. The long-run dynamics depends on the scenario considered. **a**, Without CURE, players eventually tend to adopt GTFT-like strategies. These players always reciprocate a co-player's cooperation, but they occasionally also cooperate if the co-player defected. **b**–**d**, Once CURE is available, it becomes predominant, irrespective of the tolerance threshold Δ values of 1 (**b**,**d**) or 2 (**c**), with overall payoffs remaining stable and close to the theoretical maximum (**d**). **e**,**f**, While the previous results allow for (rare) mutations, we obtain similar results if there are either no mutations at all (**e**) or if players are unable to learn CURE by mutation (**f**). For clarity, we only depict strategies that reach a frequency of at least 0.1 at some point during the process. We represent the evolving memory-1 strategies as vectors (p_1, p_2, p_3, p_4). The entries correspond to the player's cooperation probability after CC, CD, DC and DD, respectively.

also the cooperation rate of ALLC against itself. In the absence of errors, $\varepsilon \to 0$, the payoff of two CURE player thus matches the payoff of ALLC against CURE. For positive error rates, unconditional cooperators have a slight payoff advantage when they interact with a CURE opponent, compared with the interaction between two CURE players. We discuss the consequences of this advantage in more detail in the next sections.

In addition to ALLC and ALLD, we use a similar approach to derive the payoff of CURE against arbitrary memory-1 opponents. We no longer solve the respective infinitely dimensional system explicitly, but use the equations to approximate payoffs numerically (Supplementary Information Section 1 and Supplementary Tables 1–6). To validate these results, we implement independent computer simulations to estimate the players' payoffs and cooperation rates. To this end, we consider a CURE player and an opponent who adopts one of nine selected strategies. The simulation results match our analytical calculations (Supplementary Information Section 2 and Supplementary Tables 1–8). The results also suggest that, although CURE is generally cooperative, it does not cooperate with all other cooperative strategies in the presence of errors. As an example, we show that, when CURE interacts with WSLS, all four game outcomes (CC, CD, DC and DD) occur equally often over time (Supplementary Fig. 1).

Fairness and stability of CURE

The previous analysis implies that, when cumulative reciprocators interact with either ALLC or ALLD, both players obtain the same payoff. This holds more generally. We can prove that, for symmetric 2 × 2

games, CURE always enforces an equal outcome. More precisely, irrespective of the co-player's strategy σ, a cumulative reciprocator always reacts in such a way that, eventually,

$$\pi(\text{CURE}, \sigma) = \pi(\sigma, \text{CURE}).\tag{3}$$

Equation (3) holds irrespective of the precise error rate and CURE's tolerance level. Using the terminology of Duersch et al.[42,43], we conclude that CURE is unbeatable. No opponent is able to gain a lasting advantage in a direct interaction with a cumulative reciprocator. CURE shares the property of enforcing fairness with another classical strategy for the prisoner's dilemma, TFT (for which an equation like equation (3) was first derived by Press and Dyson[29]). This similarity between CURE and TFT is not a coincidence. As an illustration, consider a game without errors and suppose that Alice adopts CURE whereas Bob uses an arbitrary strategy. Then, Alice cooperates until the defection difference statistic $d(k)$ hits her tolerance level. At that point, Alice effectively implements a TFT strategy. She cooperates as long as Bob does, and she switches to defection once Bob defects. If Bob then again resumes cooperation, so does Alice.

By combining the previous results, we show that, in the absence of errors, CURE forms a Nash equilibrium in the repeated prisoner's dilemma. To see this, note that, for $\varepsilon = 0$, the payoff of CURE against itself simplifies to the mutual cooperation payoff R. If a single deviating player could achieve a larger payoff than R, equation (3) would imply that also the remaining CURE player obtains a payoff larger than R.
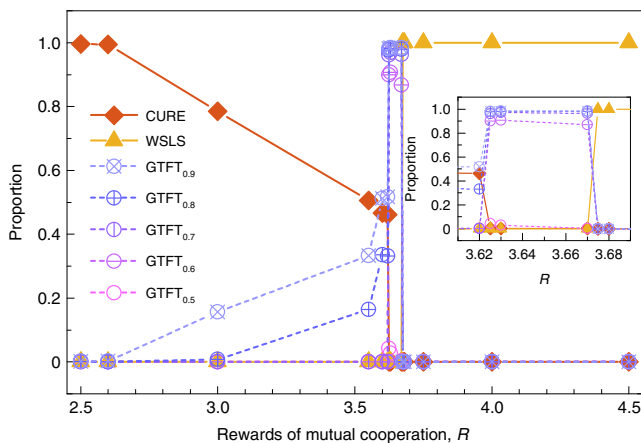
Fig. 4 | CURE, GTFT and WSLS dominate the population in different payoff regions. Our previous results are based on the classical payoffs used by Axelrod[2]. Here, we use the same basic setup as in Fig. 3 to explore which strategies are most successful as we vary the reward R for mutual cooperation. For each value of R, we record the final frequency of CURE, WSLS and the class of GTFT-like strategies of the form $(0.99, p_2, 0.99, p_4)$. We denote by GTFT_x the set of all such strategies for which p_2 and p_4 are at most x. These sets are nested, $\text{GTFT}_{0.5} \subset \text{GTFT}_{0.6} \subset \ldots \subset \text{GTFT}_{0.9}$. The graph suggests that there are two transition points, $R_1^* = 3.625$ and $R_2^* = 3.675$. When $R < R_1^*$, CURE dominates the population. When $R_1^* \leq R < R_2^*$, GTFT-like strategies are predominant. When $R \geq R_2^*$, WSLS is most abundant. For details and depictions of some of the evolutionary trajectories, see Supplementary Information Section 6.

However, because $2R$ is the maximum payoff that the two players can possibly achieve, this yields a contradiction (in the Supplementary Information, we strengthen this result slightly by showing that CURE is in fact a subgame perfect equilibrium[44]).

For positive error rates, the above argument is no longer true. Here, players can gain a payoff advantage by deviating to ALLC. However, the respective payoff advantage is often negligible. In particular, for sufficiently small error rates, CURE remains an approximate Nash equilibrium (Supplementary Information Section 1). This means that the payoff advantage from deviating to any other strategy is bounded from above, and it vanishes completely as errors become rare.

CURE and population dynamics

For the previous results, we considered games among players with fixed strategies. This kind of analysis is useful to explore a strategy's basic properties. However, it does not take into account whether players have an incentive to adopt their respective strategies in the first place. To explore this latter question, we now consider a large population of players and let their strategies evolve (see Methods for the precise setup of our evolutionary simulations).

We first examine whether CURE has an evolutionary advantage when populations contain only two strategies. We compare CURE with a tolerance level of $\Delta = 1$ with nine well-known strategies for the repeated prisoner's dilemma (Fig. 2, Supplementary Figs. 2–7 and Supplementary Information Section 3; for exact descriptions of the nine strategies, see Supplementary Information Section 2). In the simulations in Fig. 2, CURE is initially adopted by 0.1% of the population (accordingly, we speak of the other strategy as the 'resident'). The results show that CURE invades six of the nine considered resident populations (Fig. 2a–f). Against WSLS, we observe that CURE is risk-dominant (Fig. 2g)[45]. The critical frequency of cumulative reciprocators required to invade is below 50%. Only in resident populations that tend to cooperate unconditionally does CURE not evolve (Fig. 2h,i). Here, CURE suffers from its slight payoff disadvantage discussed above. However, once we additionally include defectors into the population, CURE again

becomes essential. In that case, we observe that ALLC, ALLD and CURE can stably coexist (Supplementary Figs. 8 and 9; see Supplementary Information Section 4 for a detailed analysis).

To explore the evolutionary performance of CURE when many strategies compete, we have run additional simulations for arbitrary memory-1 strategies. Each memory-1 strategy is represented by a four-dimensional vector, (p_1, p_2, p_3, p_4). Here, p_1, p_2, p_3 and p_4 refer to the player's probability to cooperate given that the outcome of the previous round is CC, CD, DC and DD, respectively. For each entry in the vector, we consider 11 possible values, equally distributed between 0.01 and 0.99 (corresponding to a noise rate of 1%). Overall, we thus allow for $11^4 = 14,641$ memory-1 strategies, to which we add a single strategy of CURE. Initially, all strategies are equally abundant. We use the same parameters and simulation techniques as above (Methods).

We examine the evolutionary dynamics under different scenarios (Fig. 3). When individuals can only choose among memory-1 strategies (not CURE), evolution eventually leads to a coexistence of different GTFT-like strategies (Fig. 3a). The respective strategies are of the form $(0.99, g_1, 0.99, g_2)$, with $g_1, g_2 \geq 0.1$. That is, all of them aim to reciprocate a co-player's cooperation, but they would occasionally also cooperate against defectors. The dynamics change completely if CURE is added. After an initial transitional period, we observe that a vast majority of players engage in CURE, irrespective of CURE's tolerance level or of the exact mutation scheme used (Fig. 3b–f). Although both GTFT and CURE are overwhelmingly cooperative, CURE seems to be more robust with respect to subsequent invasions, and it leads to slightly larger average population payoffs.

In a next step, we explore how the dynamics depends on the exact payoffs. To this end, we vary the reward for mutual cooperation R between $(T + S)/2$ and T. We identify three different regimes (Fig. 4 and Supplementary Figs. 10–17; see Supplementary Information Sections 5 and 6 for a detailed description). (1) When mutual cooperation yields substantial rewards ($R \geq 3.675$), WSLS dominates the population. This result is largely in line with earlier work. In the presence of errors, WSLS becomes evolutionarily stable once R is sufficiently large[20]. In that case, it also evolves readily among memory-1 players[13]. (2) For a small window of intermediate rewards, $3.625 \leq R < 3.675$, GTFT-like strategies are predominant. Among those strategies, most of them have a generosity of 0.6 (that is, after a co-player's defection, they cooperate with around 60% probability; Fig. 4, inset). (3) When mutual cooperation generates only a comparably modest reward ($R < 3.675$), CURE is most abundant. CURE is thus particularly likely to succeed in those environments that are traditionally considered as hostile to the evolution of cooperation.

These evolutionary results are remarkably robust with respect to various model extensions. For example, in Supplementary Information Sections 7–9, we additionally discuss the evolutionary dynamics when different variants of CURE compete, or when CURE competes against a selection of memory-2 and memory-3 strategies, or competitions between a discounted version of CURE and selected memory-1 strategies. Moreover, in Fig. 5, we illustrate how the concept of CURE extends to stochastic games[25,26] in which payoffs can fluctuate in time, and to repeated games that involve more than two players[38,39] (see also Supplementary Information Section 10).

CURE and human play

The above results highlight the strong theoretical properties of CURE. In a next step, we explore the relevance of CURE for human decision-making, by implementing a simple behavioural experiment with two treatments (Methods). In the treatment without errors, the decisions of the human participants are executed perfectly. In the treatment with errors, the decisions are mis-implemented with a 10% probability. The basic results of this experiment are as one may expect[46,47]. The game without errors yields more cooperation, and in both treatments, cooperation rates decrease slightly in time (Fig. 6a,b).
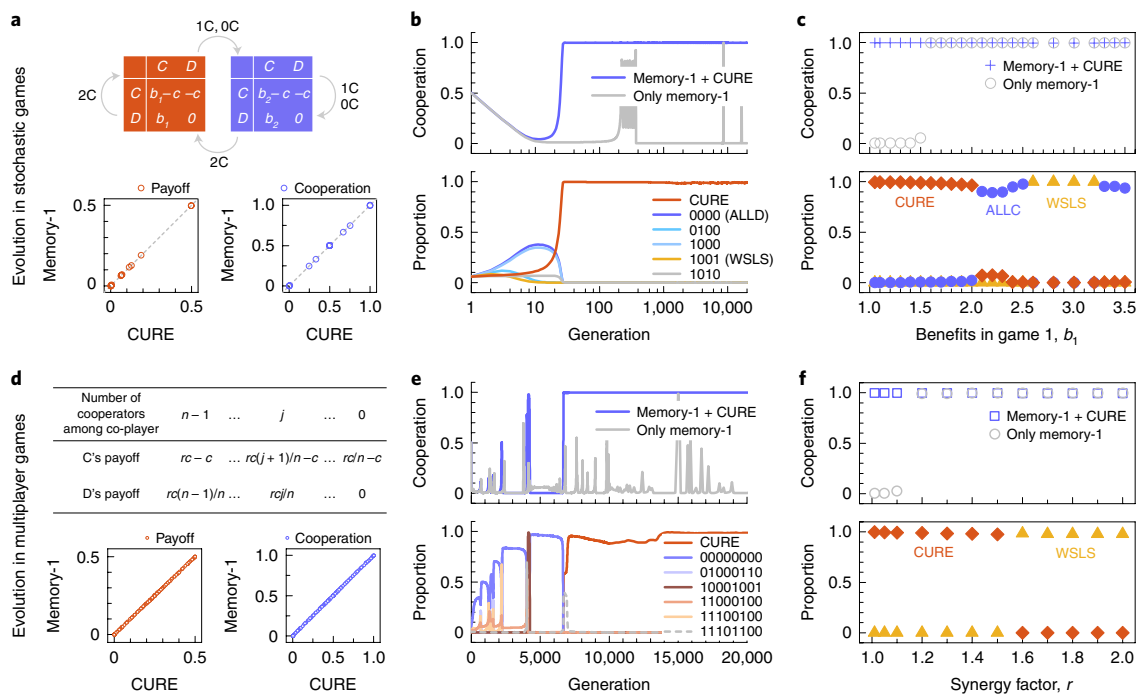
Fig. 5 | CURE in stochastic games and the repeated public goods game. In addition to the repeated prisoner's dilemma, we also explored examples in which individuals interact in a stochastic game[25,26] or in a public goods game[38,39]. **a**, As an example of a stochastic game, we suppose that players can be in one of two possible states. In both states, they interact in a donation game. The benefit of cooperation b_s depends on the current state s, with cooperation in state 1 being more valuable, $b_1 > b_2$. Players only find themselves in state 1 if they both cooperated (i.e. 2C) in the previous round. Also in this stochastic game, CURE is able to enforce equal payoffs, independent of the co-player's strategy. **b,c**, When the benefit of cooperation in the first game is comparably small, cooperation cannot evolve among memory-1 players. It evolves readily when CURE is added to

the population. **d**, As an example of a multiplayer game, we explore a public goods game with four players. Again, CURE enforces fairness. Irrespective of the strategies of the other group members, a cumulative reciprocator gets the same payoff as the other group members obtain on average. **e,f**, As before, CURE is particularly strong when cooperation is difficult to achieve otherwise (for small multiplication factors r). The parameter values in the stochastic game (unless explicitly varied) are $b_1 = 1.5$ and $b_2 = 1.01$ with a threshold of $\Delta = 1$. In the public goods game, we use $n = 4$, $r = 1.5$, $c = 1$, $\Delta = 0.5$. In both cases, the noise is $\varepsilon = 0.001$ and strategies are subject to mutations for the entire simulation. See Supplementary Information Section 10 for details on how these games were implemented and our respective implementation of CURE.

To explore the extent to which human decisions are accurately predicted by CURE, we consider one particular instance of CURE (with a threshold of $\Delta = 3$, although other thresholds yield similar results). For comparison, we consider the four pure memory-1 strategies that can sustain cooperation in the absence of errors[19], namely GRIM[32], TFT[2], firm-but-fair[27] and WSLS[12,13]. In addition, we also include three memory-2 and memory-3 strategies that have been highlighted recently: AON$_2$[20], TFT-ATFT[28] and CAPRI[21] (for definitions of all these strategies, see Supplementary Information Sections 8 and 11). For each participant in the experiment, we compare the participant's actual decisions with the decisions the participant would have made when using any of these eight strategies. For the treatment without errors, we find that all eight strategies predict human behaviour equally well. However, these results change in the treatment with errors, in which even cooperative participants defect occasionally. Here, only CURE correctly predicts the behaviour of a substantial number of participants across all rounds (Fig. 6c and Supplementary Information Section 11). Compared with traditionally considered rules of reciprocity, CURE thus seems to be a more sensible guiding principle for cooperation.

Discussion

Reciprocal cooperation requires that individuals are prepared to help others, yet they also need to be ready to fight back when their help is exploited[2]. Most of the previously known strategies with these properties, including TFT[2] and WSLS[12,13], only react to the very last round, while ignoring the entire previous history of interactions. In contrast, friends

are often comfortable with temporary imbalances, as long as they are on equal terms on average[48]. To capture reciprocal behaviours that unfold on such a longer timescale, we introduce the strategy of CURE. Individuals with this strategy keep a mental record of how often each party has defected in the past. They cooperate as long as this record is sufficiently balanced, and defect otherwise.

The strategy of CURE has a number of remarkable properties. Its payoff is robust with respect to errors, it enforces fairness and it is a Nash equilibrium in the absence of errors. Yet, it does not satisfy the notion of evolutionary stability[49]. Evolutionary stability is generally difficult to achieve in repeated games[50–52]. However, in large and heterogeneous populations, the notion of evolutionary stability seems overall less important. In such populations, the success of a strategy may rather be determined by how well it performs against a wide variety of strategies. Our simulations suggest that CURE fares particularly well in this regard. These positive results are largely independent of how often errors occur, or how frequent mutations are. In this way, CURE may revive a more general discussion on the effect of memory on the stability of reciprocity. As Press and Dyson[29] pointed out, longer memory does not give a player an immediate advantage against shorter-memory opponents. Although their assertion holds for pairwise interactions, it is no longer true when players need to find strategies that respond well to many different co-players[53]. In such mixed populations, CURE can be remarkably successful.

The analytical properties of CURE are perhaps less straightforward to derive than the properties of more conventional strategies with finite recall. For example, the dynamics among two memory-1 players

can be captured by solving a comparably simple formula[29]. In contrast, a mathematical description of the average cooperation rate of CURE leads to an infinite system of linear equations (see Supplementary Information Section 1). It is probably for this reason that related previous work either completely relied on simulations[54] or focused on a simpler case in which players always execute their actions without errors[55]. With our study, we offer a mathematical framework to analyse CURE in more general environments. Our results are applicable to classical two-player repeated games such as the prisoner's dilemma, but they also apply to multiplayer games, or games in which the feasible payoffs vary in time. Future work could explore, for example, how CURE extends to games with continuous degrees of cooperation, or it could identify alternative strategies based on cumulative assessments of the players' past history.

Even though the mathematical analysis of CURE may be elaborate, the strategy itself is straightforward to implement. Cumulative reciprocators do not need to memorize the exact outcome of the last k rounds[20], let alone the precise history of all past interactions. Instead, players only keep track of a single variable, the defection difference statistic. This statistic is easily updated, and it has an intuitive interpretation. It simply counts the imbalance in the players' past cooperative actions. A comparable bookkeeping strategy seems to be at work when people keep a loose record of obtained favours that need to be repaid. Such bookkeeping strategies may be rare among close friends, but they can play an important role in the early stages of a social relationship[56]. Especially when no mutual trust has been established, a mental record of the overall cooperation balance can be important to avoid exploitation. At the same time, it allows individuals to forgive peers who only defected by mistake. With these advantages, CURE serves as an effective mechanism to maintain fairness and mutual cooperation.

Methods

The theoretical aspects of our study raise no ethical concerns. For our behavioural experiments, we obtained institutional review board approval from the Ethics Committee of the Medical Faculty of the University of Kiel (D 613/21, 29 October 2021).

Simulation-based estimation of payoffs and cooperation rates

In addition to our analytical results and numerical approximations (Supplementary Tables 1–6), we employ simulations to estimate the players' payoffs and cooperation rates (Supplementary Tables 7 and 8). For these simulations, we conducted 1,000 independent computer experiments between any given pair of players (repeated prisoner's dilemma games) and any given group of four players (repeated public goods games). In each experiment, players interact in the game for 10,000,000 rounds. We calculate the average payoffs of the players across all rounds, and we count how often each player cooperates. The payoffs and cooperation rates are then averaged over the 1,000 experiments (over 100 experiments in the case of stochastic games). We have run these simulations for different parameter combinations. Unless noted otherwise, we implement CURE with a tolerance threshold of $\Delta = 1$. For the payoffs of the repeated prisoner's dilemma, we take the values used by Axelrod[2] as a default (that is, $T = 5$, $R = 3$, $P = 1$ and $S = 0$). Furthermore, we considered three different noise rates: low (1%), medium (5%) and high (10%). The parameters of the stochastic games and the multiplayer games are illustrated in Fig. 5. For the exact setup of the simulations used for this model extension, see Supplementary Information Section 10.

Simulating the frequency dynamics of strategies

To explore the evolution of strategy frequencies in populations of players, we consider several different scenarios:

(i) Pairwise competitions between CURE and one other strategy (Fig. 2 and Supplementary Figs. 2–7). The other strategy is either ALLC, ALLD, TFT, GTFT, contrite tit-for-tat (CTFT), WSLS, an extortionate strategy, SoftMajority or HardMajority.

(ii) A three-strategy competition between ALLC, ALLD and CURE (Supplementary Figs. 8 and 9).

(iii) Simulations of heterogeneous populations consisting of memory-1 strategies and CURE in repeated games (Figs. 3 and 4 and Supplementary Figs. 10–17).

(iv) Simulations of heterogeneous populations consisting of memory-1 strategies and CURE in stochastic games (Fig. 5b,c).

(v) Simulations of heterogeneous populations consisting of memory-1 strategies and CURE in multiplayer games (Fig. 5e,f).

(vi) Simulations in which different variants of CURE (with different tolerance levels) compete (Supplementary Fig. 18)

(vii) Simulations in which CURE competes with selected memory-2 and memory-3 strategies (Supplementary Fig. 19).

Each simulation consists of two steps. First, we obtain the payoffs $\pi(\sigma_i, \sigma_j)$ between two strategies σ_i and σ_j. The previously described simulation-based estimation is used to obtain the payoffs in most cases, including the payoff between two cumulative reciprocators, and those between a cumulative reciprocator and a player who uses either an arbitrary memory-1 strategy or one of the nine strategies selected in two-strategy competitions. For the payoffs between two memory-1 strategies, we use the analytical solution of Press and Dyson[29].

Second, we calculate the strategies' frequencies during the process through the 'survival of the fittest' in a noisy environment based on the obtained payoffs between pairs of strategies. In each generation, the evolutionary fitness of each strategy is calculated. Following Nowak and Sigmund's approach[10], the fitness of a strategy σ_i is defined by its cumulative payoff when playing the repeated game with the entire population, that is, $f(\sigma_i) = \sum_{j=1}^{n} x_j \times \pi(\sigma_i, \sigma_j)$, where n is the number of strategies in the population, and x_i and x_j are the frequencies of σ_i and σ_j, respectively. We denote the overall fitness of all strategies by $\bar{f} = \sum_{i=1}^{n} x_i f(\sigma_i)$. The frequency of σ_i in the next generation is determined to be $x_i' = x_i \times f(\sigma_i)/\bar{f}$. This elementary updating process is repeated for many generations.

To simulate evolution in two-strategy populations, each simulation ends if the frequency of each strategy no longer changes, indicating two possible steady states (that is, either the full invasion of one strategy into the other or the coexistence of the two strategies). One million generations are executed in each simulation for three-strategy populations and populations of multiple strategies. When simulating the frequency dynamics of strategies in heterogeneous populations, we sometimes allow for mutations. In that case, mutations are introduced after 2,000 generations of a simulation. The mutation rate is set to 10%. When a mutation occurs, all other strategies decrease their proportions to 99.9%, while a strategy is randomly selected to increase its proportion by 0.1%.

Behavioural experiments

The data in Fig. 6 display the results from an economic experiment. For this experiment, we recruited subjects through the online platform Prolific (www.prolific.co). In total, we report data from 172 subjects, all of whom gave their informed consent to participate. Participants were randomly allocated to one of two possible treatments. In both treatments, participants engaged in a repeated prisoner's dilemma based on our baseline payoff values (in UK pence: $R = 15$p, $S = 0$p, $T = 25$p and $P = 5$p, which correspond to the values used by Axelrod multiplied by a factor of five). Moreover, in both treatments, participants engaged for at least 20 rounds. After the 20th round, there is a constant stopping probability of 1/2 after each round, to avoid end-game effects. For better comparison, we only use data from the first 20 rounds for our statistical analysis.

The two treatments differ in how the players' intended actions are implemented. In the treatment without errors, all the players' decisions are implemented faithfully. In the treatment with errors, the players' decisions are mis-implemented with a probability of 10%. In case of an
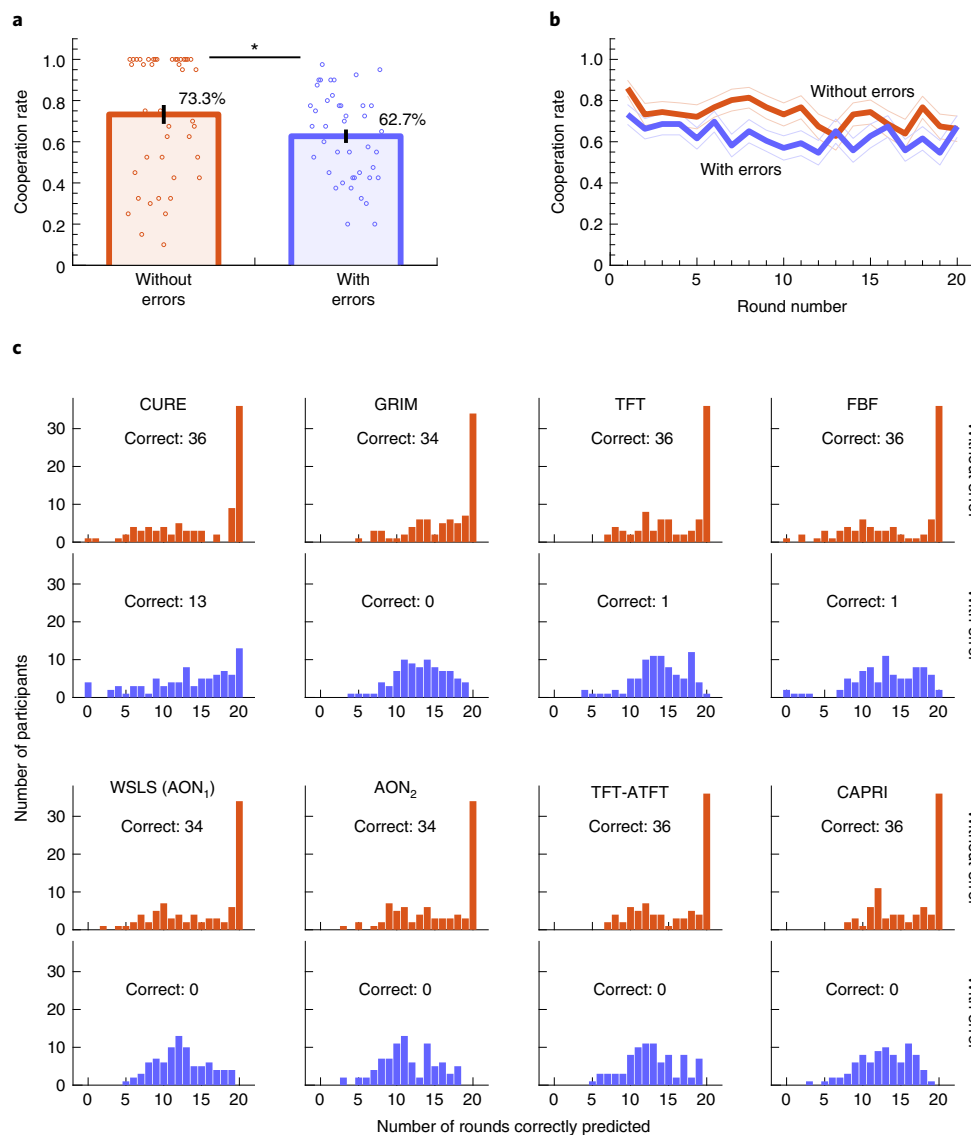
Fig. 6 | CURE in an economic experiment. To explore the relevance of CURE for human decision-making, we implemented an economic experiment based on the repeated prisoner's dilemma (see Methods for details). The experiment consists of two treatments: one treatment without errors, and one treatment with a 10% error rate. For each treatment, we report data from $n = 43$ pairs of participants. Here, we show the results with respect to the players' implemented decisions. **a**, As one may expect, there is more cooperation in the absence of errors (73.3% versus 62.7%, Mann–Whitney U test, $P = 0.018$). **b**, In both treatments, cooperation rates are slightly decreasing in time. **c**, To capture the participants' behaviour, we explore how many of their 20 decisions are explained by one of eight possible strategies. Without errors, all the considered strategies predict the participants' behaviour equally well. With errors, only CURE correctly predicts the behaviour of a substantial proportion of the subjects for all 20 rounds (this difference is significant, see Supplementary Information Section 11 for details). In **a**, dots represent the outcome of individual groups, error bars represent s.e.m. and the asterisk represents significance at $P < 0.05$. In **b**, the thick line represents averages, whereas the thin lines again represent s.e.m. All statistical tests are two-tailed and non-parametric. We do not adjust for multiple comparisons.

error, an intended cooperation is executed as a defection, and vice versa. Participants know the overall error probability, and they learn whether their own decision was mis-implemented. However, they do not know whether or not the co-player's decision was implemented faithfully.

In each round, participants learn how often each player has cooperated so far (in the treatment with errors, this number refers to how often a player's decision was implemented as cooperation). Thereafter, they make their decision of whether to cooperate in the next round. After both players have made their decision, the (implemented) outcome of that round is displayed.

To explore the extent to which human play is predicted by various well-known cooperative strategies, we consider eight possible template strategies: GRIM, tit-for-tat (TFT), firm-but-fair (FBF), WSLS, AON_2, TFT-ATFT, CAPRI and CURE (with a threshold of $\Delta = 3$). The first four strategies cooperate in the first round. Thereafter, their response is given by the memory-1 vectors

$$p_{GRIM} = (1, 0, 0, 0), \quad p_{TFT} = (1, 0, 1, 0), \quad p_{FBF} = (1, 0, 1, 1), \quad p_{WSLS} = (1, 0, 0, 1)$$

For the higher-memory strategies AON_2[20], TFT-ATFT[28] and CAPRI[21], we provide the definitions of the implemented strategies in Supplementary Information Section 11. For each participant, we compute how many of the participants' decisions are correctly predicted by each of these eight template strategies. The corresponding results are displayed in

Fig. 6c. For a more detailed description of the experimental methods, see Supplementary Information Section 11. The collected data and screenshots of our interactive game software are available online[57].

Statistics and reproducibility

For our behavioural experiment, all statistical tests are two-tailed and non-parametric. No statistical method was used to predetermine sample size. Results are based on the data of all 172 subjects who finished the experiment, not considering 15 subjects who dropped out during the instructions or 2 more subjects who dropped out during the experiment. Assignment to treatments was randomized. For further details on the study design, see Supplementary Information Section 11. The computer code for our simulations and for our behavioural experiments, as well as the resulting data, are available online[57].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Source data for Figs. 2, 3, 4 and 6 are available with this manuscript. Source data for Fig. 5 are available on Zenodo[58].

Code availability

All Java codes based on Eclipse (version Oxygen.3a release 4.7.3a), MATLAB codes (MATLAB 2020b) and oTree software (version 5.8.0b5, based on Python 3.8.10) can be obtained from Code Ocean[57].

References

1. Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971).
2. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
3. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
4. Rapoport, A., Chammah, A. M. & Orwant, C. J. *Prisoner's Dilemma: a Study in Conflict and Cooperation* (Univ. of Michigan Press, 1965).
5. Neilson, W. S. The economics of favors. *J. Econ. Behav. Organ.* **39**, 387–397 (1999).
6. Schweinfurth, M. K. & Call, J. Reciprocity: different behavioural strategies, cognitive mechanisms and psychological processes. *Learn. Behav.* **47**, 284–301 (2019).
7. Keohane, R. O. Reciprocity in international relations. *Int. Organ.* **40**, 1–27 (1986).
8. Friedman, J. W. A non-cooperative equilibrium for supergames. *Rev. Econ. Stud.* **38**, 1–12 (1971).
9. Hilbe, C., Chatterjee, K. & Nowak, M. A. Partners and rivals in direct reciprocity. *Nat. Hum. Behav.* **2**, 469–477 (2018).
10. Nowak, M. A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
11. Molander, P. The optimal level of generosity in a selfish, uncertain environment. *J. Confl. Resolut.* **29**, 611–618 (1985).
12. Kraines, D. & Kraines, V. Pavlov and the prisoner's dilemma. *Theory Decis.* **26**, 47–79 (1989).
13. Nowak, M. & Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature* **364**, 56–58 (1993).
14. Hauert, C. & Schuster, H. G. Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: a numerical approach. *Proc. R. Soc. Lond. Ser. B* **264**, 513–519 (1997).
15. Grujić, J., Cuesta, J. A. & Sanchez, A. On the coexistence of cooperators, defectors and conditional cooperators in the multiplayer iterated prisoner's dilemma. *J. Theor. Biol.* **300**, 299–308 (2012).
16. Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc. Natl Acad. Sci. U. S. A.* **110**, 15348–15353 (2013).
17. Stewart, A. J. & Plotkin, J. B. Collapse of cooperation in evolving games. *Proc. Natl Acad. Sci. U. S. A.* **111**, 17558–17563 (2014).
18. Akin, E. What you gotta know to play good in the iterated prisoner's dilemma. *Games* **6**, 175–190 (2015).
19. Akin, E. in *Ergodic Theory: Advances in Dynamical Systems* (ed. Idris Assani) 77–107 (De Gruyter, 2016).
20. Hilbe, C., Martinez-Vaquero, L. A., Chatterjee, K. & Nowak, M. A. Memory-*n* strategies of direct reciprocity. *Proc. Natl Acad. Sci. U. S. A.* **114**, 4715–4720 (2017).
21. Murase, Y. & Baek, S. K. Five rules for friendly rivalry in direct reciprocity. *Sci. Rep.* **10**, 1–9 (2020).
22. Brandt, H. & Sigmund, K. The good, the bad and the discriminator—errors in direct and indirect reciprocity. *J. Theor. Biol.* **239**, 183–194 (2006).
23. Baek, S. K., Jeong, H.-C., Hilbe, C. & Nowak, M. A. Comparing reactive and memory-one strategies of direct reciprocity. *Sci. Rep.* **6**, 1–13 (2016).
24. Hilbe, C., Wu, B., Traulsen, A. & Nowak, M. A. Cooperation and control in multiplayer social dilemmas. *Proc. Natl Acad. Sci. U. S. A.* **111**, 16425–16430 (2014).
25. Hilbe, C., Šimsa, Š., Chatterjee, K. & Nowak, M. A. Evolution of cooperation in stochastic games. *Nature* **559**, 246–249 (2018).
26. Su, Q., McAvoy, A., Wang, L. & Nowak, M. A. Evolutionary dynamics with game transitions. *Proc. Natl Acad. Sci. U. S. A.* **116**, 25398–25404 (2019).
27. Sigmund, K. *The Calculus of Selfishness* (Princeton Univ. Press, 2010).
28. Do, Yi,S., Baek, S. K. & Choi, J.-K. Combination with anti-tit-for-tat remedies problems of tit-for-tat. *J. Theor. Biol.* **412**, 1–7 (2017).
29. Press, W. H. & Dyson, F. J. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl Acad. Sci. U .S. A.* **109**, 10409–10413 (2012).
30. Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: zero-determinant strategies in noisy games. *Phys. Rev. E* **91**, 052803 (2015).
31. McAvoy, A. & Hauert, C. Autocratic strategies for iterated games with arbitrary action spaces. *Proc. Natl Acad. Sci. U. S. A.* **113**, 3573–3578 (2016).
32. Ichinose, G. & Masuda, N. Zero-determinant strategies in finitely repeated games. *J. Theor. Biol.* **438**, 61–77 (2018).
33. Reiter, J. G., Hilbe, C., Rand, D. G., Chatterjee, K. & Nowak, M. A. Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nat. Commun.* **9**, 1–8 (2018).
34. Schmid, L., Chatterjee, K., Hilbe, C. & Nowak, M. A. A unified framework of direct and indirect reciprocity. *Nat. Hum. Behav.* **5**, 1292–1302 (2021).
35. Boyd, R. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *J. Theor. Biol.* **136**, 47–56 (1989).
36. Stewart, A. J. & Plotkin, J. B. Small groups and long memories promote cooperation. *Sci. Rep.* **6**, 1–11 (2016).
37. Xia, C., Gracia-Lázaro, C. & Moreno, Y. Effect of memory, intolerance, and second-order reputation on cooperation. *Chaos* **30**, 063122 (2020).
38. Van Segbroeck, S., Pacheco, J. M., Lenaerts, T. & Santos, F. C. Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.* **108**, 158104 (2012).
39. Pinheiro, F. L., Vasconcelos, V. V., Santos, F. C., Pacheco, J. M. & Daunizeau, J. Evolution of all-or-none strategies in repeated public goods dilemmas. *PLoS Comput. Biol.* **10**, e1003945 (2014).

40. Van Veelen, M., García, J., Rand, D. G. & Nowak, M. A. Direct reciprocity in structured populations. *Proc. Natl Acad. Sci. U. S. A.* **109**, 9929–9934 (2012).

41. García, J. & van Veelen, M. In and out of equilibrium I: evolution of strategies in repeated games with discounting. *J. Econ. Theory* **161**, 161–189 (2016).

42. Duersch, P., Oechssler, J. & Schipper, B. C. Unbeatable imitation. *Games Econ. Behav.* **76**, 88–96 (2012).

43. Duersch, P., Oechssler, J. & Schipper, B. C. When is tit-for-tat unbeatable? *Int. J. Game Theory* **43**, 25–36 (2014).

44. Fudenberg, D. & Tirole, J. *Game Theory* (MIT Press, 1991).

45. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

46. Fudenberg, D., Rand, D. G. & Dreber, A. Slow to anger and fast to forgive: cooperation in an uncertain world. *Am. Econ. Rev.* **102**, 720–749 (2012).

47. Dal Bó, P. & Fréchette, G. R. Strategy choice in the infinitely repeated prisoner's dilemma. *Am. Econ. Rev.* **109**, 3929–3952 (2019).

48. Shackelford, T. K. & Buss, D. M. Betrayal in mateships, friendships, and coalitions. *Personal. Soc. Psychol. Bull.* **22**, 1151–1164 (1996).

49. Smith, J. M. & Price, G. R. The logic of animal conflict. *Nature* **246**, 15–18 (1973).

50. Boyd, R. & Lorberbaum, J. P. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* **327**, 58–59 (1987).

51. Bendor, J. & Swistak, P. Types of evolutionary stability and the problem of cooperation. *Proc. Natl Acad. Sci. U. S. A.* **92**, 3596–3600 (1995).

52. García, J. & van Veelen, M. No strategy can win in the repeated prisoner's dilemma: linking game theory and computer simulations. *Front. Robot. AI* **5**, 102 (2018).

53. Glynatsi, N. E. & Knight, V. A. Using a theory of mind to find best responses to memory-one strategies. *Sci. Rep.* **10**, 1–9 (2020).

54. Hauert, C. & Stenull, O. Simple adaptive strategy wins the prisoner's dilemma. *J. Theor. Biol.* **218**, 261–272 (2002).

55. Akin, E. Good strategies for the iterated prisoner's dilemma: Smale vs. Markov. *J. Dyn. Games* **4**, 217–253 (2017).

56. Silk, J. B. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) pp 37–54 (MIT Press, 2003).

57. Li, J. et al. Evolution of cooperation through cumulative reciprocity. *Code Ocean* https://doi.org/10.24433/CO.6894248.v1 (2022).

58. Li, J., et al. Evolution of cooperation through cumulative reciprocity. *Zenodo* https://doi.org/10.5281/zenodo.7074407 (2022).

Author contributions

H.X. conceived the original idea for this work with the help of X.Z. and designed the computational experiments, collaboratively with C.H. J.L., X.Z. and B.L. wrote the computer programs for the simulations and implemented the simulations. C.H. derived the analytical results. J.L., C.H. and H.X. analysed the results based on X.Z. and B.L.'s earlier successive contributions. C.S.L.R. designed and implemented the game software for the behavioural experiment, and collected the data. C.H. and C.S.L.R. analysed the respective results. J.L., C.H. and H.X. wrote the manuscript, and J.L. prepared the figures. H.X. and C.H. supervised the study. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Corresponding author(s):   Haoxiang Xia & Christian Hilbe

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see <u>Authors & Referees</u> and the <u>Editorial Policy Checklist</u>.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The <u>exact sample size</u> (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including <u>central tendency</u> (e.g. means) or other basic estimates (e.g. regression coefficient) AND <u>variation</u> (e.g. standard deviation) or associated <u>estimates of uncertainty</u> (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on <u>statistics for biologists</u> may be useful.

Software and code

Policy information about <u>availability of computer code</u>

Data collection	All computational simulations were performed using eclipse (Version: Oxygen.3a Release (4.7.3a)) and Matlab (R2020b). The online experiment was implemented with a game software using oTree (Version: 5.8.0b5), an open-source platform for web-based interactive tasks run in Python (Version: 3.8.10). ALL codes are publicly available in the Code Ocean (https://doi.org/10.24433/CO.6894248.v1).
Data analysis	To analyze the experimental data, we used Stata, SPSS, Matlab R2020b, and Microsoft Excel 2019 (Version: 16.64).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research <u>guidelines for submitting code & software</u> for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for Figures 2,3,4,6 is available with this manuscript. Source data for Figure 5 is available on Zenodo with DOI: 10.5281/zenodo.7074407.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	In this study we combine mathematical analyses, large-scale computational simulations, and an economic experiment. We first mathematically derived the basic properties of cumulative reciprocity (CURE). The simulation experiment further verifies these properties, and it analyzes the performance of CURE in evolving populations. Finally, we test the relevance of cumulative reciprocity for human behavior through an economic experiment. For the economic experiment, participants were matched in pairs. Each pair was randomly assigned to one of two treatments, with and without errors (Fig. 6). Each round, participants decided whether to cooperate or to defect. If both participants cooperate, each of them obtain 15 pence; if both defect, each of them obtain 5 pence; otherwise, the defector gets 25 pence and the cooperator obtains 0 pence. Without errors, the players' decisions are implemented perfectly. With errors, whenever a focal participant chooses to cooperate (defect), there is a 10% chance that the decision is mis-implemented as a defection (cooperation). Each participant received a fixed show-up fee and a variable bonus (the game payoff).
Research sample	For our computer simulations, we simulated the evolution of 2, 3, 17, 257, 14642 strategies according to the different scenarios we consider. For our economic experiment, we recruited N = 189 participants on the online platform Prolific (www.prolific.co).
Sampling strategy	For our economic experiment, we consider pairs of players as our statistical unit. Therefore, we have a statistical sample size of n = 43 for each of our two treatments.
Data collection	The simulation data was collected by running some pre-set code packages (The methods are all used in previous studies, please see the Method in the main text). For the economic experiment, we collected the data using an interface implemented with the platform oTree. The interactive task is run fully online without the presence of experimenter, who is not blind to neither the experimental condition or study hypothesis. Once the experiment is programmed and launched on the recruitment platform Prolific, data collection proceeds automatically and independently of the researcher. The researcher is only available for questions via the messaging feature of Prolific. Once data collection is complete, the data can be downloaded as a csv file for offline analysis.
Timing	The data of the economic experiment was collected on March 31, 2022, over two experimental sessions. Each session has a fixed target number of participants and stops automatically when this number is reached (about 2 hours).
Data exclusions	For the statistical analysis of our economic experiment, we have only used the first 20 rounds of groups that completed the experiment. Our results remain qualitatively unchanged if all groups and all rounds were considered. We provide a detailed drop-out analysis in the SI Section 11.
Non-participation	Fifteen participants of the economic experiment dropped out during the game's instructions. In addition, one group of 2 participants dropped out during the actual experiment (in the treatment without errors).
Randomization	For the simulations, randomization is generated by random functions in the program. For the experiment, participants were randomly matched to pairs, and each pair was randomly assigned to one of the two treatments.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	Unique biological materials
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology
☒	Animals and other organisms
☐ ☒	Human research participants

Methods

n/a	Involved in the study
☒	ChIP-seq
☒	Flow cytometry
☒	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Our study population is UK residents drawn from the online platform Prolific. The sample was made of 122 females, 49 males, 1 did not say (excluding the 17 participants who dropped out.)
Recruitment	Participants were recruited on the online platform Prolific through a standard procedure by describing the nature of this research, the length of the task, and the payoff for participating and the potential for an additional bonus payment depending on decisions made during the study. For details, see SI Section 11.