

# Cooperation in alternating interactions with memory constraints

Peter S. Park<sup>1</sup>, Martin A. Nowak<sup>1,2</sup>, Christian Hilbe<sup>3,\*</sup>

<sup>1</sup>Department of Mathematics, Harvard University, Cambridge MA 02138, USA

<sup>2</sup>Department of Organismic and Evolutionary Biology, Harvard University,  
Cambridge MA 02138, USA

<sup>3</sup>Max Planck Research Group Dynamics of Social Behavior, Max Planck Institute for Evolutionary  
Biology, 24306 Ploen, Germany

\*To whom correspondence may be addressed. Email: hilbe@evolbio.mpg.de

Cooperation is fundamental for the success of social interactions, but without mechanisms to curb free-riding it is difficult to sustain. One such mechanism is direct reciprocity, which is based on repeated interactions. An overwhelming majority of the respective research assumes synchronized decision making. In each round, individuals decide simultaneously whether to cooperate or not. Yet many manifestations of reciprocity in nature are asynchronous. Individuals provide help at one time and receive help at another. Here, we explore such alternating games in which players take turns. We mathematically characterize all Nash equilibria among the well-known memory-one strategies. The results crucially depend on how likely individuals are to make errors. If actions are executed faithfully, we prove that the same strategies that sustain cooperation in the simultaneous game also apply to the alternating game. However, once individual actions are subject to noise, the results are different. In the simultaneous game, deterministic strategies like Win-Stay Lose-Shift can maintain cooperation. In contrast, the alternating game requires individuals to randomize. After a defection of the opponent, they need to forgive with a probability that depends on the previous outcome, the error rate, and the benefit to cost ratio. We probe these conclusions with various model extensions, exploring the effect of discounted games, irregular alternation patterns, and higher memory. In all cases, we observe that mutual cooperation still evolves for a wide range of parameter values. However, compared to simultaneous games, alternating games require different strategies to cope with noise, and none of the respective strategies are evolutionarily stable.

Cooperation can be maintained by direct reciprocity, where individuals help others in repeated interactions<sup>1–3</sup>. Traditionally, researchers capture the logic of direct reciprocity with the repeated prisoner’s dilemma<sup>4–14</sup>. According to that model, two individuals – usually referred to as players – interact with each other over several rounds. In each round, both players can either cooperate or defect. Mutual cooperation yields a better payoff than mutual defection, but each individual has an incentive to defect. Theoretical and experimental work suggests that cooperation can evolve if there are sufficiently many interactions between the individuals<sup>15</sup>. This work has been used to explain a wide variety of behaviors, including why humans are more likely to cooperate in stable groups<sup>16</sup>, why certain animal species share food<sup>17</sup>, and why firms are able to achieve higher market prices when they engage in collusion<sup>18</sup>.

A standard assumption that underlies much of this research is that individuals make their decisions simultaneously (or at least in ignorance of the co-player’s current decision). We refer to this kind of repeated interaction as a *simultaneous game* (**Fig. 1a**). For many natural manifestations of reciprocity, however, simultaneous cooperative exchanges are unlikely or even impossible, such as when people ask for favors<sup>19</sup>, vampire bats donate blood to their conspecifics<sup>17</sup>, sticklebacks engage in predator inspection<sup>20</sup>, or ibis take turns when leading their flock<sup>21</sup>. Such sequential interactions are better captured by *alternating games*, in which players consecutively decide whether to cooperate<sup>22–25</sup>. When individuals decide asynchronously, they make their decisions based on different histories. The most recent events one player has in memory differ from the most recent events that the next player takes into account (**Fig. 1b**). Such asymmetries in turn make it more difficult to successfully coordinate on cooperation. As a result, many well-known strategies like Tit-for-Tat or Win-Stay Lose-Shift fail to evolve when players move alternately<sup>22,23</sup>. Instead, previous computational<sup>22–24</sup> and experimental studies<sup>26</sup> suggest that individuals need to be more forgiving. However, a full understanding of optimal play in alternating games is lacking, even though optimal behavior in the simultaneous game is by now well-analyzed<sup>27–32</sup>.

Here, we propose an analytical approach to describe when cooperation evolves in the alternating game. In line with the previous literature, we typically focus on individuals with so-called *memory-one strategies*<sup>3</sup>. Memory-one strategies depend on each player’s most recent move. Our analysis involves two steps. First, we show that successful play in alternating games does not require a sophisticated cognitive apparatus. More specifically, when interacting with a given memory-one opponent, it suffices to respond with a *reactive strategy* that only depends on the co-player’s most recent move. This result is reminiscent of a previous finding of Press and Dyson for the simultaneous game<sup>33</sup>. They showed that against a memory-one strategy, there is nothing to gain from having a *larger* memory than the opponent. Our result for the alternating game goes one step further. Against a memory-one strategy, players can afford to have a strictly *lower* memory, without any loss to their or their co-player’s payoff. As we show, this result crucially depends on the alternating move structure; it is not true when players move simultaneously. In a second step, we show that in order to identify a best response to a given memory-one player, we only need to check the four most extreme reactive strategies: unconditional defection, unconditional cooperation, Tit-for-Tat and Anti-Tit-for-Tat. Using this approach, we identify all Nash equilibria among the memory-one strategies.

In the absence of errors, we find an unexpected equivalence. The very same memory-1 strategies that can be used to enforce cooperation in the simultaneous game also enforce cooperation in the alternating game. However, once we take into account errors, the predictions for the two models diverge. In the simultaneous game, Win-Stay Lose-Shift is evolutionarily stable when the benefit to cost ratio is sufficiently large and when errors are sufficiently rare<sup>34,35</sup>. In that case, there is a simple rule for how to sustain full cooperation: individuals should repeat their previous action if it yielded a sufficiently large payoff, and switch to the opposite action otherwise. In contrast, in the alternating game, all stable cooperative strategies require players to randomize. After mutual defection, they need to cooperate with some well-defined probability that depends on the game parameters and the error rate. Although the respective strategies in the alternating game are Nash equilibria, we show that none of them is evolutionarily stable. As a result, evolving cooperation rates in the alternating game often tend to be lower than in the simultaneous game, although this difference is smaller than perhaps expected from static stability considerations alone. We summarize our analytical findings in **Table 1**.

Our work suggests that in most realistic scenarios, successful play in alternating games requires different kinds of behaviors than predicted by the earlier theory on simultaneous games. In this way, we corroborate earlier experimental work on human cooperation<sup>26</sup>, and provide new theoretical methods to analyze repeated games in the future. Overall, we find that cooperation is still feasible in alternating games. However, the strategies that enforce cooperation can be neutrally invaded, and hence cooperation tends to be more short-lived than in the simultaneous game.

## Model

**Model description.** In the following, we formulate a simple baseline scenario, which we use to derive our main analytical results. More general scenarios are discussed in a later section, and in full detail in the **Supplementary Information**. We consider interactions between two players, player 1 and player 2. Both players repeatedly decide whether to cooperate ( $C$ ) or defect ( $D$ ). These repeated interactions can take place in two different ways. In the *simultaneous game*, there is a discrete number of rounds. In each round, both players make their decision at the same time, not knowing their co-player's decision (**Fig. 1a**). In contrast, in the *alternating game*, the players move consecutively. We consider the strictly alternating game: Player 1 moves first, and then player 2 learns about player 1's decision and moves next (**Fig. 1b**). We note that there are also variants of the alternating game in which the order of moves is random<sup>22,25</sup>. In particular, one player may by chance make two or more consecutive moves before it is the other player's turn again. The effect of such irregular alternation patterns will be discussed later.

For the simultaneous game, the possible payoffs in each round can be represented by four parameters. Players receive the reward  $R$  in rounds in which they both cooperate; they receive the temptation payoff  $T$  and the sucker's payoff  $S$ , respectively, if only one player cooperates; and they receive the punishment payoff  $P$  in case they both defect. For  $T > R > P > S$ , we obtain the prisoner's dilemma. In the alternating game, however, it is useful to assume that payoffs can be assigned to each player's individual

action<sup>22</sup>. In that case, the value of one player's cooperation is independent of the co-player's previous or subsequent decision (or equivalently, payoffs are independent of how the two players' decisions are grouped into rounds). As a result, we obtain the donation game<sup>3</sup>. Here, cooperation means to pay a cost  $c > 0$  in order to provide a benefit  $b > c$  to the co-player. The donation game is a special case of a prisoner's dilemma for which

$$R = b - c, \quad S = -c, \quad T = b, \quad P = 0. \quad (1)$$

To compare the alternating game with the simultaneous game, we assume payoffs satisfy (1) throughout.

In the baseline scenario, we consider infinitely repeated games, and we study players who make their decisions based on each player's most recent move. In the simultaneous game, the respective strategies are called *memory-1 strategies*<sup>36</sup>; they take into account the outcome of one previous round (**Fig. 1a**). Such strategies can be represented as a 4-tuple,  $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ . The entry  $p_{ij}$  denotes the probability to cooperate in the next round. This probability depends on the player's action  $i$  and the co-player's action  $j$  in the previous round. The equivalent strategy class also exists in alternating games<sup>22</sup>. In alternating games, however, there is no longer a unique previous round to which both players refer. Instead, the last round that is taken into account depends on the perspective of each player. It consists of the respective last moves of the two players (**Fig. 1b**). An important subset of memory-1 strategies are the so-called reactive strategies. Here, players ignore their own previous action, and only condition their behavior on what the co-player previously did. Reactive strategies are therefore those memory-1 strategies for which  $p_{CC} = p_{DC}$  and  $p_{CD} = p_{DD}$ .

Some well-known examples of memory-1 strategies for the simultaneous game include Always Defect,  $ALLD = (0, 0, 0, 0)$ , Tit-for-Tat,  $TFT = (1, 0, 1, 0)$ , and Win-Stay Lose-Shift,  $WSLS = (1, 0, 0, 1)$ . In the alternating game, a strategy called *Firm-but-Fair*<sup>3</sup>, defined by  $FBF = (1, 0, 1, 1)$  and also referred to as *Forgiver*<sup>24</sup>, has been successful in evolutionary competitions. Out of these examples,  $ALLD$  and  $TFT$  are reactive, whereas  $WSLS$  and  $FBF$  are not. We say a strategy is deterministic if each conditional cooperation probability is either zero or one. In particular, all of the above examples are deterministic. Otherwise, we call the strategy stochastic.

Note that our analysis includes the possibility that players sometimes make errors. That is, when a player decides to cooperate, there is some probability  $\varepsilon$  that the player defects by mistake. Conversely, a player who intends to defect may cooperate with the same probability. We refer to the case of  $\varepsilon = 0$  as the game without errors, and to  $\varepsilon > 0$  as the game with errors. We note that even a strategy that is deterministic becomes fully stochastic in the game with errors, because in that case a player's effective cooperation probability is always between  $\varepsilon$  and  $1 - \varepsilon$ .

Considering memory-1 strategies is useful for two reasons. First, such strategies are straightforward to interpret, and the respective conditional probabilities can be easily inferred from experiments<sup>26</sup>. Second, when both players use memory-1 strategies, their average payoffs are simple to compute (see also **Methods**). To this end, suppose player 1 uses the strategy  $\mathbf{p}$  and player 2 adopts strategy  $\mathbf{q}$ . By representing the game as a Markov chain, we can compute the stationary distribution  $\mathbf{v} = (v_{CC}, v_{CD}, v_{DC}, v_{DD})$ .

The entries of this stationary distribution give the probabilities of observing each of the four possible combinations of the players' actions over the course of the game. Based on this stationary distribution, we define player 1's payoff as  $\pi(\mathbf{p}, \mathbf{q}) = (v_{CC} + v_{DC})b - (v_{CC} + v_{CD})c$ , and similarly for player 2. While the baseline scenario focuses on memory-1 strategies, our results are more general. For example, when we describe which memory-1 strategies are Nash equilibria in the following, co-players are allowed to deviate to strategies with arbitrarily long (but finite) memory. Moreover, similar approaches can also be used to explore the evolutionary dynamics of memory-2 strategies, as we will discuss later.

## Results

**A recipe for identifying Nash equilibria for alternating games.** To predict which memory-1 strategies evolve in the alternating game, we first characterize which of them are Nash equilibria. In the following, we refer to a strategy  $\mathbf{q}$  as a Nash equilibrium if  $\pi(\mathbf{q}, \mathbf{q}) \geq \pi(\mathbf{p}, \mathbf{q})$  for all alternative memory-1 strategies  $\mathbf{p}$  (for stronger results, see **Supplementary Information**). That is, against a co-player who adopts the Nash equilibrium strategy  $\mathbf{q}$ , a player has no incentive to choose any different memory-1 strategy. The notion of Nash equilibrium is closely related to evolutionary robustness<sup>27</sup>. In a population of size  $N$ , a resident strategy  $\mathbf{q}$  is called evolutionary robust if no mutant strategy  $\mathbf{p}$  has a fixation probability larger than neutral,  $1/N$ . When selection is sufficiently strong, strategies are evolutionary robust if and only if they are Nash equilibria<sup>28</sup>.

Verifying that a given strategy  $\mathbf{q}$  is a Nash equilibrium is not straightforward. In principle, this requires us to compare its payoff to the payoff of all possible mutant strategies  $\mathbf{p}$ , taken from the uncountable set of all memory-1 strategies. However, for alternating games, it is possible to simplify the task in two steps (see **Supplementary Information** for details). The first step is to show that it is sufficient to compare  $\mathbf{q}$  to all reactive strategies, a strategy set of lower dimension. The intuition for this result is as follows. Even if player 1 starts out with an arbitrary memory-1 strategy  $\mathbf{p}$ , it is always possible to find an associated reactive strategy  $\tilde{\mathbf{p}}$  that yields the same stationary distribution and the same payoff against  $\mathbf{q}$  (**Fig. 2**). That is, to find a best response to a strategy that remembers both players' last moves, it is sufficient to explore all strategies that only remember the co-player's last move. In particular, not only is there no advantage of having a strictly larger memory than the opponent, as shown by Press and Dyson for simultaneous games<sup>33</sup>. A player can afford to remember strictly less in the alternating game.

The second step is to show that we do not need to consider all reactive strategies to find a best response against  $\mathbf{q}$ . Instead, it suffices to consider all deterministic reactive strategies. By combining these two steps, it becomes straightforward to check whether a given memory-1 strategy is a Nash equilibrium. We only need to compare its payoff against itself to the four payoffs that can be achieved by deviating to Always Defect (*ALLD*), Always Cooperate (*ALLC*), Tit-for-Tat (*TFT*), or Anti-Tit-for-Tat (*ATFT*).

**Equilibria in alternating games without errors.** Using the above recipe, we first explore which memory-1 strategies can sustain full cooperation in games without errors (see **Supplementary Infor-**

**mation** for all derivations). To this end, we call a memory-1 strategy a *partner*<sup>15,29</sup> if (i) it is fully cooperative against itself, and (ii) if it is a Nash equilibrium (such strategies are referred to as ‘good’ by Akin<sup>30,37,38</sup>). We find that partners are exactly those memory-1 strategies  $\mathbf{q}$  for which the following three conditions hold,

$$q_{CC} = 1, \quad q_{CD} \leq 1 - \frac{c}{b-c}q_{DD}, \quad q_{CD} \leq 1 - \frac{c}{b}q_{DC}. \quad (2)$$

The first condition is needed to ensure that the strategy is fully cooperative against itself. The other two conditions restrict how cooperative a player is allowed to be after having been exploited by the co-player. If these last two conditions are violated, the strategy  $\mathbf{q}$  can either be invaded by *ALLD* or *ATFT*. Together, the three requirements in (2) define a three-dimensional polyhedron within the space of all memory-1 strategies (**Fig. 3a**). The volume of this polyhedron increases with the benefit to cost ratio  $b/c$ . While the polyhedron never contains *ALLC*, it always contains the conditionally cooperative strategies *TFT* and *GRIM* (for these two strategies, we additionally require the respective players to cooperate in the very first round to ensure payoffs are well-defined, see **Supplementary Information**). Moreover, for  $b \geq 2c$ , the polyhedron contains *WSLS* and *FBF* (independent of the outcome of the first round).

Similarly, we can also identify all Nash equilibria where the players mutually defect. We refer to the respective strategies as *defectors*. We obtain the following necessary and sufficient condition,

$$q_{DD} = 0, \quad q_{DC} \leq \frac{c}{b}(1 - q_{CD}), \quad q_{DC} \leq \frac{c}{b-c}(1 - q_{CC}). \quad (3)$$

Again, the first equation ensures that two players with the respective strategy end up mutually defecting against each other. The other two conditions ensure that the strategy is comparably unresponsive towards a co-player who tries to initiate cooperation. Similar to before, the three conditions define a three-dimensional polyhedron (**Supplementary Fig. 2a**). The set of defectors is non-empty for all parameter values, and it always contains the strategy *ALLD*.

Finally, we identify a third class of Nash equilibria, referred to as *equalizers*<sup>39</sup>. As in the simultaneous game<sup>33</sup>, equalizers are strategies that unilaterally control the co-player’s payoff. If one player adopts an equalizer strategy, the co-player’s payoff is fixed, independent of the co-player’s strategy<sup>40–44</sup>. In the alternating game, these strategies are characterized by

$$q_{CD} = \frac{b q_{CC} - c(1 + q_{DD})}{b - c}, \quad q_{DC} = \frac{b q_{DD} + c(1 - q_{CC})}{b - c}. \quad (4)$$

When both players adopt an equalizer strategy, neither player has anything to gain from deviating; the resulting outcome is a Nash equilibrium.

We also show a converse result: If a memory-1 strategy for the alternating game is a Nash equilibrium, then it either needs to be a partner, a defector, or an equalizer. Remarkably, the same three strategy classes also arise as Nash equilibria of the simultaneous game<sup>28</sup>. Even the algebraic conditions for being a partner, defector, or equalizer coincide (however, the existing proof for the simultaneous game<sup>28</sup> is

somewhat more intricate than the proof for the alternating game that we provide in the **Supplementary Information**). There is, however, one difference. In the simultaneous game, there is a fourth class of Nash equilibria, referred to as ‘alternators’<sup>28</sup>. Alternators cooperate in one round, only to defect in the next. In the **Supplementary Information**, we show that such patterns of behavior cannot emerge among memory-1 players in the alternating game.

**Equilibria in alternating games with errors.** Next, we explore how the Nash equilibria change when we introduce errors. In the following, we discuss the case of partner strategies; the analogous results for defectors and equalizers are derived in the **Supplementary Information**. For partner strategies, we find that errors impose additional constraints. First, partners only exist when errors are sufficiently rare,  $\varepsilon < \frac{1}{2} \left(1 - \frac{c}{b}\right)$ . Second, the respective conditions are now considerably more restrictive,

$$q_{CC} = q_{DC} = 1, \quad q_{CD} \leq 1 - \frac{c}{(1-2\varepsilon)b}, \quad q_{DD} = \frac{(1-2\varepsilon)(b+\varepsilon c q_{CD}) - c}{(1-2\varepsilon)(b+\varepsilon c)}. \quad (5)$$

In particular, if the co-player cooperated in the previous round, partners are strictly required to cooperate in the next round, independent of their own previous action (because now  $p_{DC} = 1$ ). If the co-player defected, partners need to cooperate with a well-defined probability, as defined by the last two conditions in (5). The last condition guarantees that neither *ALLC* nor *TFT* has a selective advantage against *q*. In the game without errors, this requirement is satisfied automatically. There, all strategies with  $q_{CC} = 1$  yield the full cooperation payoff  $b-c$  against each other. In the game with errors, however, such strategies are no longer neutral. Instead, they differ in how quickly they are able to restore cooperation after an error, and to which extent they are able to capitalize on their co-players’ mistakes. Noisy environments thus impose additional constraints on self-cooperative strategies to be stable.

As a result of these additional constraints, the three-dimensional polyhedron degenerates to a one-dimensional line segment (**Fig. 3b**). On one end of this line segment, there is Generous Tit-for-Tat, which also arises in the simultaneous game<sup>6,45</sup>,

$$GTFT = \left(1, 1 - \frac{c}{(1-2\varepsilon)b}, 1, 1 - \frac{c}{(1-2\varepsilon)b}\right) \quad (6)$$

On the other end of this line segment, we find a strategy that resembles the main characteristics of Firm-but-Fair<sup>3</sup>; we thus refer to this strategy as *Stochastic Firm But Fair*,

$$SFBF = \left(1, 0, 1, \frac{(1-2\varepsilon)b-c}{(1-2\varepsilon)(b+\varepsilon c)}\right) \quad (7)$$

Behaviors similar to *SFBF* have been observed in early simulations of alternating games<sup>22,23</sup>. There, it was found that evolutionary trajectories often lead to strategies that are deterministic, except that they randomize after mutual defection. Our results provide an analytical justification: *SFBF* is the only such strategy that is a Nash equilibrium.

The above conditions in (5) provide a complete characterization of all partner strategies in the alternating game with errors. Despite decades of research, an analogous characterization for the simultaneous game is not yet available (**Table 1**). However, it is known that particular strategies, most importantly *WSLS*, can be evolutionarily stable in the presence of noise<sup>34</sup>. That is, in the simultaneous game, cooperation can be sustained with a simple deterministic strategy if  $b > 2c$ . In contrast, conditions (5) imply that no such deterministic strategy is available in the alternating game. Moreover, while the partner strategies characterized by (5) are Nash equilibria, we show in the **Supplementary Information** that they all are vulnerable to neutral invasion by either *ALLC* or *TFT* (in fact to all strategies with  $q_{CC} = q_{DC} = 1$ ). These results suggest that cooperation can still evolve in alternating games, but it may be less robust than in the simultaneous game.

**Evolutionary dynamics of alternating games.** In order to test these equilibrium predictions, we next explore which behaviors emerge when the players’ strategies are subject to evolution. To this end, we consider a population of  $N$  players. Each member of the population is equipped with a memory-1 strategy. They obtain payoffs by interacting with all other population members. To model the spread of successful strategies, we assume individuals with high payoffs are imitated more often<sup>46</sup> (or equivalently, such individuals produce more offspring<sup>47</sup>). In addition, new strategies are introduced through random exploration (or equivalently, through mutations). These random strategies are uniformly taken from the space of all memory-1 strategies. We capture the resulting dynamics with computer simulations. For details, see **Methods**.

First, we explore the evolutionary dynamics for fixed game parameters. We record which strategies the players use over the course of evolution to sustain cooperation. In **Fig. 3**, we represent those strategies that yield a cooperation rate against themselves of at least 80%; other threshold values lead to similar conclusions. We call these strategies ‘self-cooperative’. By definition, players with these strategies are likely to cooperate after mutual cooperation. Here we are thus interested in how they react when either one or both players defected. Without errors, the respective conditional cooperation probabilities show quite some variation. As a result, the distributions in **Fig. 3c** are comparably flat. Overall, players act in such a way that the partner conditions (2) are satisfied, but they show no preference for a particular partner strategy. Once we allow for errors, the evolving strategies change (**Fig. 3d**). Players tend to always cooperate if the co-player did so in the previous round, with  $q_{CC} \approx q_{DC} \approx 1$ . Moreover, after mutual defection, they cooperate with some strictly positive probability. Both patterns are predicted by our equilibrium conditions (5). We find a similar match between static theory and evolutionary simulations for defectors, or when we explore evolution in the simultaneous game (**Figures S1–S3**).

In a next step, we compare the dynamics of the alternating and the simultaneous game across different parameter values. To this end, we systematically vary the benefit of cooperation, the population size, the selection strength, and the mutation rate (**Fig. 4**). In games without errors, we observe hardly any difference between the alternating and the repeated game. Both games yield almost identical cooperation rates over time, and these cooperation rates are similarly affected by parameter changes. A difference



between the two games only becomes apparent when players need to cope with errors. Here, the simultaneous game leads to systematically higher cooperation rates than the alternating game. This difference is most visible for intermediate benefit-to-cost ratios and intermediate error rates, as one may expect: For small benefits and frequent errors, cooperation evolves in neither game, whereas for large benefits and rare errors, cooperation evolves in both games (**Supplementary Fig. 4**).

**Evolutionary results beyond the baseline scenario.** Our baseline scenario represents an idealized model of alternating interactions. It assumes (i) the game is infinitely repeated, (ii) players move in a strictly alternating fashion, (iii) games take place in a well-mixed population, and (iv) players use memory-1 strategies. In the following, we use simulations to explore the effect of each of these assumptions in turn. Here, we briefly summarize the respective results. For an exact description of the models, and for a more detailed discussion of the results, we refer to the **Supplementary Information**.

We start by considering games with finitely many rounds. To incorporate a finite game length, we assume that each time both players have made a decision, the game continues with a constant probability  $\delta$ . **Fig. 5a-c** shows the respective evolutionary results for  $\delta = 0.96$  (such that games last for 25 rounds on average). We observe similar results as in the infinitely repeated game: The simultaneous game leads to more cooperation (**Fig. 5a**); moreover, if players cooperate, their strategies exhibit the main characteristics of *WSLS* in the simultaneous game, and of *SFBF* and *GTFT* in the alternating game (**Fig. 5b**). Further simulations suggest that these qualitative results hold when players interact for at least 10 rounds (**Supplementary Fig. 5**). When interactions are shorter, cooperation is unlikely to evolve at all (**Fig. 5c**).

In a next step, we explore irregular alternation patterns. To this end, we assume that every time a player has made a decision, with probability  $s$  it is the other player who moves next. We refer to  $s$  as the game’s switching probability. For  $s = 1$ , we recover the baseline scenario, in which players strictly alternate. For  $s = 1/2$ , the player to move next is determined randomly. Simulations suggest that in both cases, players again use strategies akin to *GTFT* and *SFBF* to sustain cooperation (**Fig. 5e**). However, the robustness of the strategies depends on the switching probability. In particular, mutual cooperation is most likely to evolve when players alternate regularly (**Fig. 5f**, **Supplementary Fig. 6**).

To explore the effect of population structure, we follow the framework of Brauchli et al<sup>48</sup>. Instead of well-mixed populations, players are now arranged on a two-dimensional lattice. They use memory-1 strategies to engage in pairwise interactions with each of their neighbors. For the simultaneous game, we recover the main results of Brauchli et al<sup>48</sup>: population structure can further enhance cooperation, and it makes it more likely that strategies similar to *WSLS* evolve (**Fig. 5g-i**). For the alternating game, we observe that cooperation remains the most abundant outcome, but spatial structure does not necessarily result in homogeneous populations any longer. Instead, in some simulations we find cooperative and non-cooperative strategies to stably coexist (one particular instance is shown in **Fig. 5h**).

Finally, we also analyzed the impact of larger memory. Exploring the dynamics among general memory- $k$  strategies is not straightforward, as the strategy space increases rapidly. For instance, while there are only 16 pure memory-1 strategies, there are 65,536 memory-2 strategies and more than  $10^{19}$

memory-3 strategies<sup>35</sup>. We thus confine ourselves to pure memory-2 strategies in the following. In a first step, we explored which of these strategies are evolutionarily stable, see **Supplementary Fig. 7a**. For the simultaneous game, we find many such strategies, including several strategies with high cooperation rates. In the alternating game, we only find one generic evolutionarily stable strategy, *ALLD*. Nevertheless, with respect to the evolving cooperation rates, stochastic evolutionary trajectories hardly show any difference between alternating and simultaneous games. The two games differ, however, in terms of the strategies that evolve, and in how robust these strategies are (**Supplementary Fig. 7b-e**).

## Discussion

An overwhelming majority of past research on reciprocity deals with repeated games where individuals simultaneously decide whether to cooperate<sup>3,15</sup>. In contrast, most natural occurrences of reciprocity require asynchronous acts of giving. Cooperation routinely takes the form of assisting a peer, providing a gift, or taking the lead in a joint endeavor<sup>19–21</sup>. In such examples, simultaneous cooperation can be unfeasible, undesirable, or unnecessary. Herein, we have thus explored which strategies arise in alternating games where individuals make their decisions in turns. In such games, one individual's cooperation is reciprocated not immediately, but at some point in the future.

To explore the dynamics of cooperation in alternating games, we first describe all Nash equilibria among the memory-1 strategies. Memory-1 strategies are classical tools that have been used to describe the evolutionary dynamics of repeated games for several decades<sup>22,36,49</sup>. However, most of the early work on memory-1 strategies was restricted to evolutionary simulations. Only with the pioneering work of Press and Dyson<sup>33</sup> and others<sup>27–32</sup>, better mathematical techniques have become available. Using these techniques, it has become possible to describe all Nash equilibria of the infinitely repeated simultaneous game without errors<sup>28</sup>. Herein, we make similar progress for the alternating game, both for the case with and without errors (for the simultaneous game with errors, a complete characterization of the Nash equilibria remains an open problem, see **Table 1**).

Our results suggest that there are both unexpected parallels and important differences between simultaneous and alternating games. The parallels arise when individuals do not make errors. Here, the two models of reciprocity make the same predictions about the feasibility of cooperation. Cooperation evolves in the same environments, and it can be maintained using the same strategies. However, once individuals make mistakes, the predictions of the two models diverge. First, the two models require different kinds of strategies to maintain cooperation. In the simultaneous game, cooperation can be sustained with the deterministic memory-1 strategy Win-Stay Lose-Shift<sup>36,49</sup>. Individuals with that strategy simply reiterate their previous behavior if it was successful, and they switch their behavior otherwise. In contrast, in the alternating game, no simple deterministic rules for cooperation exist. Although there are still infinitely many memory-1 strategies that can maintain cooperation, all of them require individuals to randomize occasionally. One example of such a strategy for alternating games is Stochastic Firm-But-Fair (*SFBF*). Individuals with this strategy always reciprocate a co-player's cooperation, never tolerate

exploitation, and cooperate with some intermediate probability if both players defected. Similar behaviors have been observed in earlier simulations<sup>22,23</sup>. Our results provide a theoretical underpinning: *SFBF* is the unique memory-1 strategy that can sustain cooperation while retaliating against unconditional defectors in the strongest possible way.

The simultaneous game and the alternating game also differ in how stable cooperation is in evolving populations. In the simultaneous game, the evolution of cooperation is hardly affected by errors, provided the error rate is below a certain threshold (**Fig. 4, Supplementary Fig. 4**). In some instances, errors can even enhance cooperation<sup>50</sup>. This body of work is based on the insight that evolutionarily stable cooperation is impossible in simultaneous games without errors<sup>51–55</sup>. For any cooperative resident, it is always possible to find neutral mutant strategies that eventually lead to the demise of cooperation. However, once individuals occasionally commit errors, a strategy like *WSLS* is no longer neutral with respect to other cooperative strategies; it becomes evolutionarily stable<sup>34,52</sup>. The situation is different in alternating games. Even in the presence of rare errors, strategies like *SFBF* remain vulnerable. They can be invaded by unconditional cooperators or by any other strategy that fully reciprocates a co-player's cooperation.

Despite these differences in the stability of their main strategies, evolving cooperation rates in the simultaneous and the alternating game are often surprisingly similar. To interpret these results, we note that when evolution is stochastic and takes place in finite populations, no strategy persists indefinitely. Even evolutionarily stable strategies are invaded eventually. As a result, the overall abundance of cooperation is not only determined by the stability of any given strategy. Instead, it depends on additional aspects, such as the time it takes cooperative strategies to reappear when they are invaded. The relative importance of these different aspects depends on the details of the considered evolutionary process. To further illustrate these observations, we have run additional simulations for memory-1 players with local mutations<sup>56</sup> (see **Supplementary Information**). Because evolutionary stability considerations are less relevant when mutations are local, we observe that the cooperation rates of the alternating and the simultaneous game become more similar (**Supplementary Fig. 8**).

Cooperation is defined as a behavior where individuals pay a cost in order to increase the payoff or fitness of someone else<sup>2</sup>. When individuals interact repeatedly, such cooperative interactions can be maintained by reciprocity. Here we have argued that in many examples, reciprocity arises as a series of asynchronous acts of cooperation. Most often, people do favors not to be rewarded immediately, but to request similar favors in future. Such consecutive acts of cooperation also appear to be at work when vampire bats<sup>17</sup>, sticklebacks<sup>20</sup>, ibis<sup>21</sup>, tree swallows<sup>57</sup>, or macaques<sup>58</sup> engage in reciprocity. We have shown that mutual cooperation is still possible in such alternating exchanges. But compared to the predominant model of reciprocity in simultaneous games, cooperation requires different kinds of strategies, and it is more volatile.

## Methods

**Calculation of payoffs.** When two players with memory-1 strategies interact, their expected payoffs can be computed by representing the game as a Markov chain<sup>3</sup>. To this end, suppose the first player's strategy is  $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ , and the second player's strategy is  $\mathbf{q} = (q_{CC}, q_{CD}, q_{DC}, q_{DD})$ . Depending on the most recent actions of the two players (which can be either  $CC$ ,  $CD$ ,  $DC$ , or  $DD$ ), we can compute how likely we are to observe each of the four outcomes in the following round. For the alternating game, we obtain the following transition matrix<sup>22</sup>,

$$M_A = \begin{pmatrix} p_{CC}q_{CC} & p_{CC}(1-q_{CC}) & (1-p_{CC})q_{CD} & (1-p_{CC})(1-q_{CD}) \\ p_{CD}q_{DC} & p_{CD}(1-q_{DC}) & (1-p_{CD})q_{DD} & (1-p_{CD})(1-q_{DD}) \\ p_{DC}q_{CC} & p_{DC}(1-q_{CC}) & (1-p_{DC})q_{CD} & (1-p_{DC})(1-q_{CD}) \\ p_{DD}q_{DC} & p_{DD}(1-q_{DC}) & (1-p_{DD})q_{DD} & (1-p_{DD})(1-q_{DD}) \end{pmatrix}. \quad (8)$$

Based on this transition matrix, we compute how often players observe each of the four outcomes. To this end, we solve the equation for the stationary distribution,  $\mathbf{v} = \mathbf{v}M_A$ . In most cases, the solution of this equation is unique. Uniqueness is guaranteed, for example, when the players' strategies  $\mathbf{p}$  and  $\mathbf{q}$  are fully stochastic, or when the error rate is positive. In exceptional cases, however, the transition matrix can allow for two or more stationary distributions. In that case, the outcome of the game is still well-defined, after specifying how players act in the very first round.

Given the stationary distribution  $\mathbf{v} = (v_{CC}, v_{CD}, v_{DC}, v_{DD})$ , we define the players' payoffs as

$$\begin{aligned} \pi_1 &= (v_{CC} + v_{DC})b - (v_{CC} + v_{CD})c, \\ \pi_2 &= (v_{CC} + v_{CD})b - (v_{CC} + v_{DC})c. \end{aligned} \quad (9)$$

This definition implicitly assumes that the game is indefinitely repeated and that future payoffs are not discounted. However, analogous formulas can be given in case there is a constant continuation probability  $\delta$ , or equivalently, if future payoffs are discounted by  $\delta$  (see **Supplementary Information**).

We compare our results for the alternating game with the corresponding results for the standard repeated prisoner's dilemma, where players decide simultaneously. Payoffs for the simultaneous game can be calculated in the same way as before. Only the transition matrix needs to be replaced by<sup>3</sup>

$$M_S = \begin{pmatrix} p_{CC}q_{CC} & p_{CC}(1-q_{CC}) & (1-p_{CC})q_{CC} & (1-p_{CC})(1-q_{CC}) \\ p_{CD}q_{DC} & p_{CD}(1-q_{DC}) & (1-p_{CD})q_{DC} & (1-p_{CD})(1-q_{DC}) \\ p_{DC}q_{CD} & p_{DC}(1-q_{CD}) & (1-p_{DC})q_{CD} & (1-p_{DC})(1-q_{CD}) \\ p_{DD}q_{DD} & p_{DD}(1-q_{DD}) & (1-p_{DD})q_{DD} & (1-p_{DD})(1-q_{DD}) \end{pmatrix}. \quad (10)$$

Although the two matrices share many similarities, the resulting dynamics can be very different. For example, if the two players use *TFT*, then the matrix  $M_S$  allows for three invariant sets (corresponding to mutual cooperation, mutual defection, and alternating cooperation). However, the respective matrix  $M_A$

only allows for the first two invariant sets<sup>22</sup>. More generally,  $M_S$  allows for equilibria where players cooperate in one round but defect in the next round. Such equilibria are impossible for  $M_A$  (see **Supplementary Information**).

We sometimes assume players commit errors. We incorporate errors by assuming that with probability  $\varepsilon$ , a player who intends to cooperate defects by mistake. Analogously, a player who wishes to defect cooperates instead with the same probability. Such errors are straightforward to incorporate into the model. For  $\varepsilon > 0$ , a player's strategy  $\mathbf{p}$  translates into an effective strategy  $\mathbf{p}^\varepsilon := (1 - \varepsilon)\mathbf{p} + \varepsilon(\mathbf{1} - \mathbf{p})$ . To compute the payoffs of strategy  $\mathbf{p}$  against strategy  $\mathbf{q}$  in the presence of errors, we apply the formulas (8) – (10) to the strategies  $\mathbf{p}^\varepsilon$  and  $\mathbf{q}^\varepsilon$ .

**Evolutionary dynamics.** In the following, we describe the evolutionary process for the baseline scenario. For the various model extensions (**Fig. 5**, **Supplementary Fig. 5**– **Supplementary Fig. 8**), we use appropriately adapted versions of this process, as described in more detail in the **Supplementary Information**. To model how successful strategies spread in well-mixed populations, we use a pairwise comparison process<sup>46</sup>. This process considers a population of constant size  $N$ . Initially, all population members are unconditional defectors. Each player derives a payoff by interacting with all other population members; for each pairwise interaction, payoffs are given by Eq. (9).

To model how strategies with a high payoff spread within a population, we consider a model in discrete time. In each time step, one player is chosen from the population at random. This player is then given an opportunity to revise its strategy. The player can do so in two ways. First, with probability  $\mu$  (the mutation rate), the player may engage in random strategy exploration. In this case, the player discards its strategy, and samples a new strategy uniformly at random from the set of all memory-1 strategies. Second, with probability  $1 - \mu$ , the player considers imitating one of its peers. In this case, the player selects a random role model from the population. If the role model's payoff is  $\pi_R$  and the focal player's payoff is  $\pi_F$ , then imitation occurs with a probability given by the Fermi function<sup>59</sup>

$$\rho = \frac{1}{1 + \exp \left[ -\beta(\pi_R - \pi_F) \right]}. \quad (11)$$

If imitation occurs, the focal player discards its previous strategy and adopts the role model's strategy instead. In the formula for the imitation probability, the parameter  $\beta \geq 0$  is called the strength of selection. It measures the extent to which players are guided by payoff differences between the players' strategies. For  $\beta = 0$ , any payoff differences are irrelevant. The focal player adopts the role model's strategy with probability  $1/2$ . As  $\beta$  becomes larger, payoff differences become increasingly important. In the limiting case  $\beta \rightarrow \infty$ , imitation only occurs if the role model's payoff at least matches the focal player's payoff.

Overall, the two mechanisms of random strategy exploration and directed strategy imitation give rise to a stochastic process on the space of all population compositions. For positive mutation rates, this process is ergodic. In particular, the average cooperation rate (as a function of the number of time steps) converges, and it is independent of the considered initial population. Herein, we have explored

this process with computer simulations. We have recorded which strategies the players adopt over time and how often they cooperate on average. For most of these simulations, we assume that mutations are sufficiently rare<sup>60</sup>. For those simulations, we require mutant strategies to either fix in the population or to go extinct before the next mutation occurs. Under this regime, the mutant’s fixation probability can be computed explicitly<sup>9</sup>. This in turn allows us to simulate the evolutionary dynamics more efficiently<sup>61,62</sup>.

**Parameters and specific procedures used for the figures.** For the simulations in well-mixed populations, we used the following baseline parameters:

$$\begin{aligned}
\text{Benefit of cooperation: } & b=3 \\
\text{Cost of cooperation: } & c=1 \\
\text{Population size: } & N=100 \\
\text{Selection strength: } & \beta=5 \text{ (Fig. 3, Figs. S1–S3) and } \beta=1 \text{ (all other figures)} \\
\text{Error rate: } & \varepsilon=0 \text{ (without errors), or } \varepsilon=0.02 \text{ (with errors)} \\
\text{Mutation rate: } & \mu \rightarrow 0.
\end{aligned} \tag{12}$$

Changes in these parameters are systematically explored in **Fig. 4** and **Supplementary Fig. 4**. For **Fig. 3**, **Fig. 4**, and **Supplementary Fig. 1–Supplementary Fig. 6**, the respective simulations are run for at least  $10^7$  time steps each (measured in number of introduced mutant strategies over the course of a simulation). For **Fig. 5**, **Supplementary Fig. 7**, and **Supplementary Fig. 8**, simulations are run for a shorter time (as illustrated in the respective panels that illustrate the resulting dynamics); however, here all results are obtained by averaging over 50 – 200 independent simulations.

To report which strategies the players use to sustain cooperation (or defection), we record all strategies that arise during a simulation that have a cooperation rate against themselves of at least 80% (in the case of self-cooperators), or a cooperation rate of less than 20% (in the case of self-defectors). In **Fig. 3**, **Supplementary Fig. 1–Supplementary Fig. 3**, and **Supplementary Fig. 5**, we show the marginal distributions of all strategies that we have obtained in this way. For these distributions, each strategy is weighted by how long the strategy has been present in the population. In **Fig. 5**, **Supplementary Fig. 7**, and **Supplementary Fig. 8**, we represent the self-cooperative strategies by computing the average of the respective marginal distributions. In some cases (**Fig. 5e**, **Supplementary Fig. 7**, **Supplementary Fig. 8**), we also report how robust self-cooperative strategies are on average. To this end, we record for each self-cooperative resident strategy how many mutants need to be introduced into the population until a mutant strategy reaches fixation. We consider self-cooperative strategies that resist invasion by many mutant strategies as more robust.

Finally, for the simulations for spatial populations (**Fig. 5g–i**), we closely follow the setup of Brauchli et al<sup>48</sup>. Here, we consider a population of size  $N=2,500$ . Players are arranged on a  $50 \times 50$  lattice with periodic boundary conditions. Players use memory-1 strategies (initially they adopt the strategy *ALLD*). In every generation, every player interacts in a pairwise game with each of its eight immediate neighbors. After these interactions, all players are independently given an opportunity to update their strategies.

With probability  $\mu = 0.002$ , an updating player chooses a random strategy, uniformly taken from all memory-1 strategies (global mutations). With probability  $1 - \mu$ , the updating player adopts the strategy of the neighbor with the highest payoff (but only if this neighbor's payoff is better than the focal player's payoff). This elementary process is then repeated for 20,000 generations. Panels **Fig. 5g,i** show averages across 50 independent simulations of the process. Panel **Fig. 5h** illustrates two particular realizations.

**Code availability.** All simulations were performed with MATLAB\_R2019b. The respective code is available at: [https://osf.io/v5hgd/?view\\_only=de839809032b498394ff462363bc75b2](https://osf.io/v5hgd/?view_only=de839809032b498394ff462363bc75b2).

**Data availability.** The raw data generated with these computer simulations is also available at: [https://osf.io/v5hgd/?view\\_only=de839809032b498394ff462363bc75b2](https://osf.io/v5hgd/?view_only=de839809032b498394ff462363bc75b2).

## References

- [1] Trivers, R. L. The evolution of reciprocal altruism. *The Quarterly Review of Biology* **46**, 35–57 (1971).
- [2] Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
- [3] Sigmund, K. *The Calculus of Selfishness* (Princeton Univ. Press, Princeton, NJ, 2010).
- [4] Rapoport, A. & Chammah, A. M. *Prisoner's Dilemma* (University of Michigan Press, Ann Arbor, 1965).
- [5] Axelrod, R. *The evolution of cooperation* (Basic Books, New York, NY, 1984).
- [6] Nowak, M. A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
- [7] Szabó, G., Antal, T., Szabó, P. & Droz, M. Spatial evolutionary prisoner's dilemma game with three strategies and external constraints. *Physical Review E* **62**, 1095–1103 (2000).
- [8] Killingback, T. & Doebeli, M. The continuous Prisoner's Dilemma and the evolution of cooperation through reciprocal altruism with variable investment. *The American Naturalist* **160**, 421–438 (2002).
- [9] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).
- [10] Doebeli, M. & Hauert, C. Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology Letters* **8**, 748–766 (2005).
- [11] Szolnoki, A., Perc, M. & Szabó, G. Phase diagrams for three-strategy evolutionary prisoner's dilemma games on regular graphs. *Physical Review E* **80**, 056104 (2009).

- [12] García, J. & Traulsen, A. The structure of mutations and the evolution of cooperation. *PLoS One* **7**, e35287 (2012).
- [13] Grujic, J. *et al.* A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Scientific Reports* **4**, 4615 (2014).
- [14] Glynatsi, N. & Knight, V. A bibliometric study of research topics, collaboration and centrality in the field of the Iterated Prisoner's Dilemma. *Humanities and Social Sciences Communications* **8**, 45 (2021).
- [15] Hilbe, C., Chatterjee, K. & Nowak, M. A. Partners and rivals in direct reciprocity. *Nature Human Behaviour* **2**, 469–477 (2018).
- [16] Melis, A. P. & Semmann, D. How is human cooperation different? *Philosophical Transactions of the Royal Society B* **365**, 2663–2674 (2010).
- [17] Wilkinson, G. S. Reciprocal food-sharing in the vampire bat. *Nature* **308**, 181–184 (1984).
- [18] Bernheim, D. & Whinston, M. D. Multimarket contact and collusive behavior. *The RAND journal of economics* **21**, 1–26 (1990).
- [19] Jackson, M. O., Rodriguez-Barraquer, T. & Tan, X. Social capital and social quilts: Network patterns of favor exchange. *American Economic Review* 1857–976 (2012).
- [20] Milinski, M. Tit For Tat in sticklebacks and the evolution of cooperation. *Nature* **325**, 433–435 (1987).
- [21] Voelkl, B. *et al.* Matching times of leading and following suggest cooperation through direct reciprocity during V-formation flight in ibis. *Proceedings of the National Academy of Sciences USA* **112**, 2115–2120 (2015).
- [22] Nowak, M. A. & Sigmund, K. The alternating prisoner's dilemma. *Journal of Theoretical Biology* **168**, 219–226 (1994).
- [23] Frean, M. R. The prisoner's dilemma without synchrony. *Proceedings of the Royal Society B* **257**, 75–79 (1994).
- [24] Zagorsky, B. M., Reiter, J. G., Chatterjee, K. & Nowak, M. A. Forgiver triumphs in alternating prisoner's dilemma. *PLoS One* **8**, e80814 (2013).
- [25] McAvoy, A. & Hauert, C. Autocratic strategies for alternating games. *Theoretical Population Biology* **113**, 13–22 (2016).
- [26] Wedekind, C. & Milinski, M. Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proceedings of the National Academy of Sciences USA* **93**, 2686–2689 (1996).
- [27] Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences USA* **110**, 15348–15353 (2013).
- [28] Stewart, A. J. & Plotkin, J. B. Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences USA* **111**, 17558 – 17563 (2014).
- [29] Hilbe, C., Traulsen, A. & Sigmund, K. Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games and Economic Behavior* **92**, 41–52 (2015).



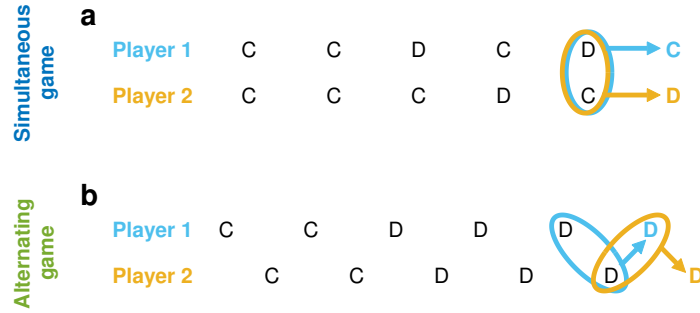
- [30] Akin, E. What you gotta know to play good in the iterated prisoner’s dilemma. *Games* **6**, 175–190 (2015).
- [31] McAvoy, A. & Nowak, M. A. Reactive learning strategies for iterated games. *Proceedings of the Royal Society A* **475**, 20180819 (2019).
- [32] Glynatsi, N. & Knight, V. Using a theory of mind to find best responses to memory-one strategies. *Scientific Reports* **10**, 1–9 (2020).
- [33] Press, W. H. & Dyson, F. D. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. *PNAS* **109**, 10409–10413 (2012).
- [34] Lorberbaum, J. P., Bohning, D. E., Shastri, A. & Sine, L. E. Are there really no evolutionarily stable strategies in the iterated prisoner’s dilemma? *Journal of Theoretical Biology* **214**, 155–169 (2002).
- [35] Hilbe, C., Martinez-Vaquero, L. A., Chatterjee, K. & Nowak, M. A. Memory- $n$  strategies of direct reciprocity. *Proceedings of the National Academy of Sciences USA* **114**, 4715–4720 (2017).
- [36] Nowak, M. A. & 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).
- [37] Akin, E. The iterated prisoner’s dilemma: Good strategies and their dynamics. In Assani, I. (ed.) *Ergodic Theory, Advances in Dynamics*, 77–107 (de Gruyter, Berlin, 2016).
- [38] Akin, E. Good strategies for the iterated prisoner’s dilemma: Smale vs. Markov. *Journal of Dynamics and Games* **4**, 217–253 (2017).
- [39] Boerlijst, M. C., Nowak, M. A. & Sigmund, K. Equal pay for all prisoners. *American Mathematical Monthly* **104**, 303–307 (1997).
- [40] Hilbe, C., Nowak, M. A. & Sigmund, K. The evolution of extortion in iterated prisoner’s dilemma games. *Proceedings of the National Academy of Sciences USA* **110**, 6913–6918 (2013).
- [41] Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: Zero-determinant strategies in noisy games. *Physical Review E* **91**, 052803 (2015).
- [42] McAvoy, A. & Hauert, C. Autocratic strategies for iterated games with arbitrary action spaces. *Proceedings of the National Academy of Sciences* **113**, 3573–3578 (2016).
- [43] Ichinose, G. & Masuda, N. Zero-determinant strategies in finitely repeated games. *Journal of Theoretical Biology* **438**, 61–77 (2018).
- [44] Mamiya, A. & Ichinose, G. Zero-determinant strategies under observation errors in repeated games. *Physical Review E* **102**, 032115 (2020).
- [45] Molander, P. The optimal level of generosity in a selfish, uncertain environment. *Journal of Conflict Resolution* **29**, 611–618 (1985).
- [46] Traulsen, A., Pacheco, J. M. & Nowak, M. A. Pairwise comparison and selection temperature in evolutionary game dynamics. *Journal of Theoretical Biology* **246**, 522–529 (2007).
- [47] Wu, B., Bauer, B., Galla, T. & Traulsen, A. Fitness-based models and pairwise comparison models of evolutionary games are typically different—even in unstructured populations. *New Journal of Physics* **17**, 023043 (2015).

- [48] Brauchli, K., Killingback, T. & Doebeli, M. Evolution of cooperation in spatially structured populations. *Journal of Theoretical Biology* **200**, 405–417 (1999).
- [49] Kraines, D. P. & Kraines, V. Y. Pavlov and the prisoner’s dilemma. *Theory and Decision* **26**, 47–79 (1989).
- [50] Zhang, H. Errors can increase cooperation in finite populations. *Games and Economic Behavior* **107**, 203–219 (2018).
- [51] Boyd, R. & Lorberbaum, J. No pure strategy is evolutionary stable in the iterated prisoner’s dilemma game. *Nature* **327**, 58–59 (1987).
- [52] Boyd, R. Mistakes allow evolutionary stability in the repeated Prisoner’s Dilemma game. *Journal of Theoretical Biology* **136**, 47–56 (1989).
- [53] Lorberbaum, J. M. D. No strategy is evolutionary stable in the repeated Prisoner’s Dilemma. *Journal of Theoretical Biology* **168**, 117–130 (1994).
- [54] García, J. & van Veelen, M. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *Journal of Economic Theory* **161**, 161–189 (2016).
- [55] García, J. & van Veelen, M. No strategy can win in the repeated prisoner’s dilemma: Linking game theory and computer simulations. *Frontiers in Robotics and AI* **5**, 102 (2018).
- [56] Stewart, A. J. & Plotkin, J. B. The evolvability of cooperation under local and non-local mutations. *Games* **6**, 231–250 (2015).
- [57] Lombardo, M. P. Mutual restraint in tree swallows: A test of the tit for tat model of reciprocity. *Science* **227**, 1363–1365 (1985).
- [58] Muroyama, Y. Mutual reciprocity of grooming in female japanese macaques (*macaca fuscata*). *Behaviour* **119**, 161–170 (1991).
- [59] Szabó, G. & Tóke, C. Evolutionary Prisoner’s Dilemma game on a square lattice. *Physical Review E* **58**, 69–73 (1998).
- [60] Wu, B., Gokhale, C. S., Wang, L. & Traulsen, A. How small are small mutation rates? *Journal of Mathematical Biology* **64**, 803–827 (2012).
- [61] Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory* **131**, 251–262 (2006).
- [62] Imhof, L. A. & Nowak, M. A. Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B* **277**, 463–468 (2010).
- [63] Szolnoki, A. & Perc, M. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Scientific Reports* **4**, 5496 (2014).
- [64] Szolnoki, A. & Perc, M. Evolution of extortion in structured populations. *Physical Review E* **89**, 022804 (2014).
- [65] Wu, Z.-X. & Rong, Z. Boosting cooperation by involving extortion in spatial prisoner’s dilemma games. *Physical Review E* **90**, 062102 (2014).
- [66] Hilbe, C., Hagel, K. & Milinski, M. Asymmetric power boosts extortion in an economic experiment. *PLoS ONE* **11**, e0163867 (2016).

**Acknowledgments.** P.S.P. is supported by the National Science Foundation through the Graduate Research Fellowship Program (grant number DGE1745303), by the Centre for Effective Altruism through the Global Priorities Fellowship, and by Harvard University through graduate student fellowships. M.A.N. is supported by the Army Research Laboratory (grant W911NF-18-2-0265), the Bill & Melinda Gates Foundation (grant OPP1148627), the John Templeton Foundation (grant 61443), and the Office of Naval Research (grant N00014-16-1-2914). C.H. acknowledges generous funding by the Max Planck Society and by the ERC Starting Grant E-DIRECT (850529).

**Author contributions.** P.S.P., M.A.N., and C.H: Designed the research; P.S.P., M.A.N., and C.H: Performed the research; P.S.P., M.A.N., and C.H: Wrote the paper.

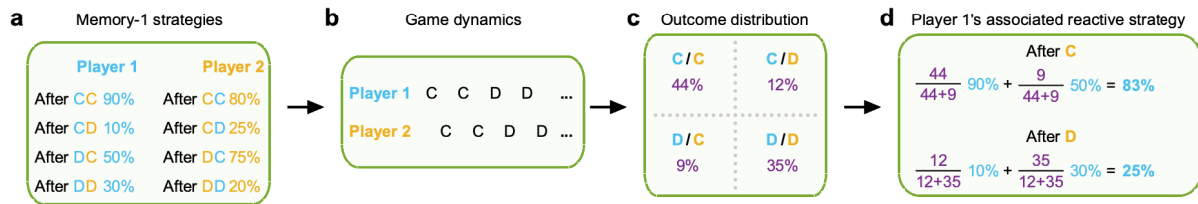
**Competing interests.** The authors declare no conflict of interest.



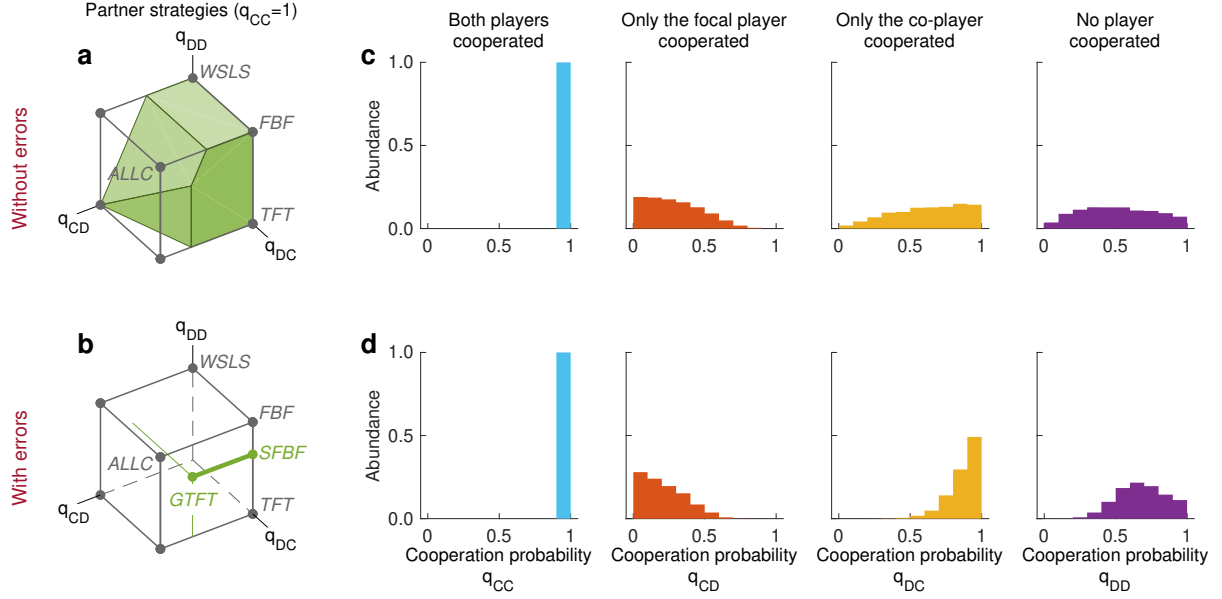
**Figure 1: Game dynamics for the simultaneous and the alternating game.** In both the simultaneous and the alternating game, two players interact repeatedly. In each turn, they decide whether to cooperate ( $C$ ) or to defect ( $D$ ). In the simultaneous game, they make their decision at the same time (or at least not knowing the other player's decision). In the alternating game, one player decides before the other player does. In both cases, we study memory-1 strategies. In that case, an individual's next action only depends on each individual's previous action. We illustrate the information each individual takes into account for its last decision with colored ellipses. In the simultaneous game, individuals take into account the same information. In the alternating game, decisions are based on different sets of information.

	Simultaneous game	Alternating game
Game without errors	<p><b>a</b></p> <ul style="list-style-type: none"> <li>The full set of partner strategies is <b>known</b>.</li> <li>Partner strategies exist for all <math>b &gt; c</math>.</li> <li>A strategy is a partner if and only if <math display="block">q_{CC} = 1,</math> <math display="block">q_{CD} \leq 1 - \frac{c}{b-c} q_{DD},</math> <math display="block">q_{CD} \leq 1 - \frac{c}{b} q_{DC}.</math> </li> <li>None of these strategies is evolutionarily stable.</li> </ul>	<p><b>b</b></p> <ul style="list-style-type: none"> <li>We characterize partner strategies <b>here</b> (Eq. [2]).</li> <li>Partner strategies exist for all <math>b &gt; c</math>.</li> <li>A strategy is a partner if and only if <math display="block">q_{CC} = 1,</math> <math display="block">q_{CD} \leq 1 - \frac{c}{b-c} q_{DD},</math> <math display="block">q_{CD} \leq 1 - \frac{c}{b} q_{DC}.</math> </li> <li>None of these strategies is evolutionarily stable.</li> </ul>
Game with errors	<p><b>c</b></p> <ul style="list-style-type: none"> <li>The full set of partner strategies is <b>not known</b>.</li> <li>Two examples are known: <math display="block">GTFT \text{ if } b &gt; c \text{ and } \epsilon &lt; \frac{1}{2} \left(1 - \frac{c}{b}\right)</math> <math display="block">WSLS \text{ if } b \geq 2c \text{ and } \epsilon \leq \frac{1}{2} \left(1 - \frac{c}{\sqrt{(b-c)c}}\right)</math> </li> <li><i>WSLS</i> is evolutionarily stable if the respective inequalities are strict.</li> </ul>	<p><b>d</b></p> <ul style="list-style-type: none"> <li>We characterize partner strategies <b>here</b> (Eq. [5]).</li> <li>Partner strategies exist when <math>b &gt; c</math> and <math>\epsilon &lt; \frac{1}{2} \left(1 - \frac{c}{b}\right)</math>.</li> <li>A strategy is a partner if and only if <math display="block">q_{CC} = q_{DC} = 1,</math> <math display="block">q_{CD} \leq 1 - \frac{c}{(1-2\epsilon)b}, \quad q_{DD} = \frac{(1-2\epsilon)(b+\epsilon c q_{CD})-c}{(1-2\epsilon)(b+\epsilon c)}.</math> </li> <li>None of these strategies is evolutionarily stable.</li> </ul>

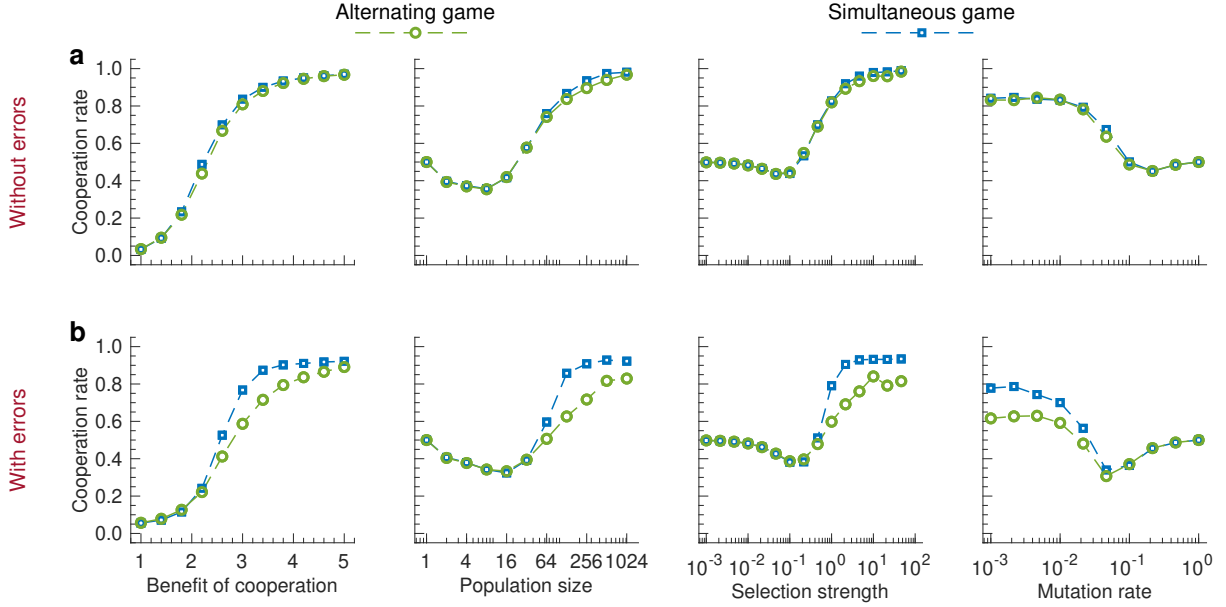
**Table 1: A characterization of partners among the memory-1 strategies.** Within the class of memory-1 strategies, we provide an overview of the strategies that sustain full cooperation in a Nash equilibrium. The respective strategies are called partner strategies, or partners<sup>15</sup>. **a**, For the simultaneous game without errors, partners have been first described by Akin<sup>30,37</sup> (he calls them ‘good strategies’). Akin’s approach has been extended by Stewart and Plotkin<sup>28</sup> to describe all memory-1 Nash equilibria of the simultaneous game. In the absence of errors, none of these strategies is evolutionarily stable<sup>51,53</sup>. Instead, one can always find neutral mutant strategies which act as a stepping stone out of equilibrium<sup>54</sup>. **b**, For the alternating game without errors, Eq. (2) provides a full characterization of all partner strategies. Cooperation is maintained by the same strategies as in the simultaneous game. **c**, Despite decades of research, the exact set of partner strategies for the simultaneous game with errors is not known. However, there are at least two instances of partner strategies, *GTFT*<sup>6,45</sup>, and Win-Stay Lose-Shift, *WSLS*<sup>36,49</sup>. For repeated games with errors, evolutionary stability is generally feasible<sup>52</sup>. In particular, *WSLS* is evolutionarily stable if the benefit to cost ratio is sufficiently large and if errors are sufficiently rare<sup>34</sup>. **d**, For the alternating game with errors, we characterize all partner strategies in Eq. (5). None of them is deterministic. As a result, none of them is evolutionarily stable (see **Supplementary Information** for details).



**Figure 2: In alternating games, individuals can afford to remember less than their opponent.** We prove the following result: if two memory-1 players interact, any of the players can switch to a simpler reactive strategy (that only depends on the co-player's previous action) without changing the resulting payoffs. Here, we illustrate this result for player 1. **a**, Initially, both players use memory-1 strategies. That is, a player's cooperation probability depends on the most recent decision of each player. There are four conditional cooperation probabilities. **b**, The strategies determine how players interact in the alternating game. **c**, Based on the strategies, we can compute how often we are to observe each pairwise outcome over the course of the game by calculating the game's stationary distribution. **d**, Based on the stationary distribution, and on player 1's memory-1 strategy, we can compute an associated reactive strategy. This reactive strategy only consists of two conditional cooperation probabilities. They determine what to do if the co-player cooperated (or defected) in the previous round. The cooperation probabilities can be calculated as a weighted average of the respective memory-1 strategy's cooperation probabilities. The resulting reactive strategy for player 1 yields the same outcome distribution against player 2 as the original memory-1 strategy. We note that for this result, the assumption of alternating moves is crucial. In the simultaneous game, the respectively defined reactive strategy does not yield the same outcome distribution against player 2 as the original memory-1 strategy (see **Supplementary Information**).

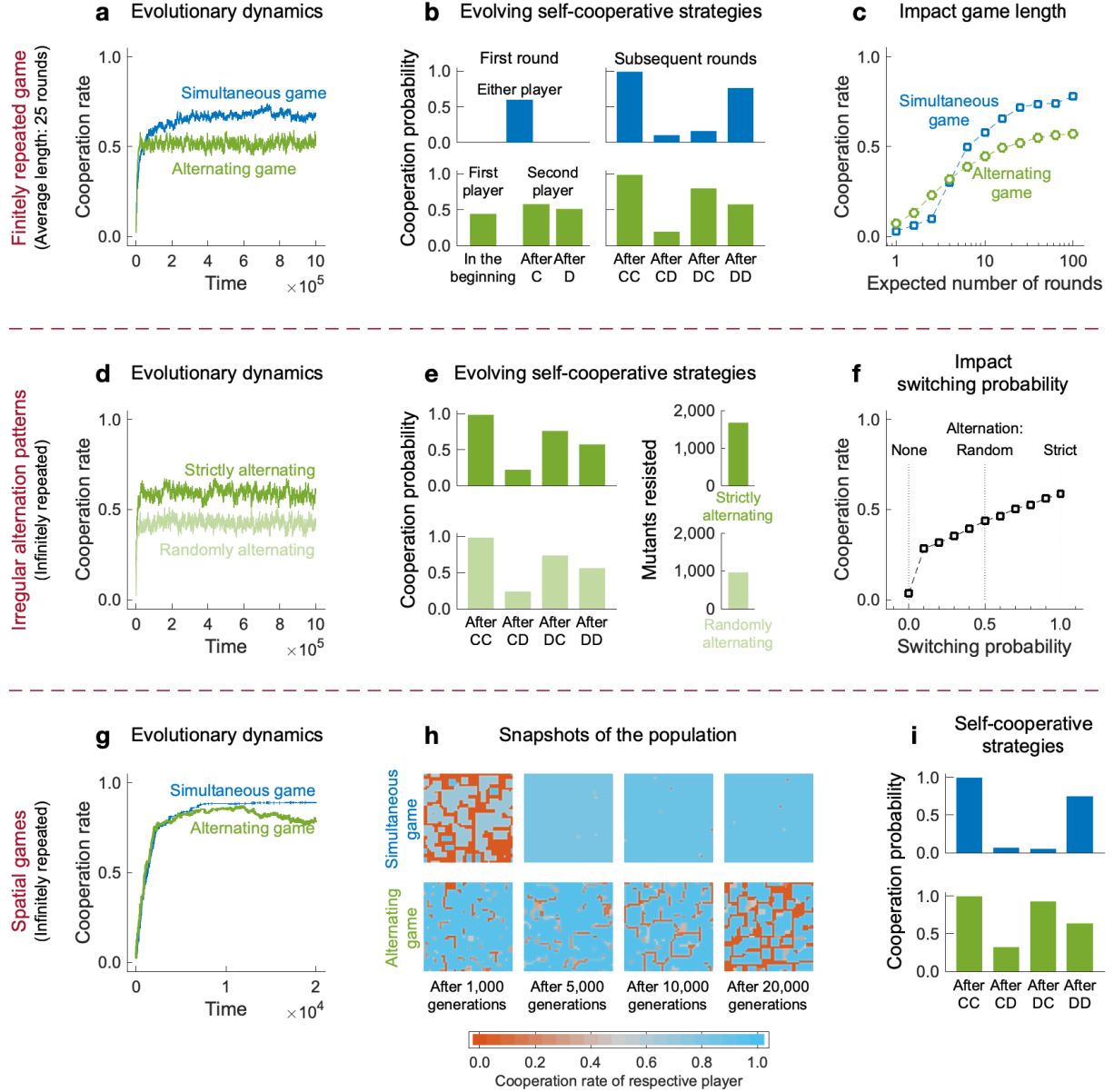


**Figure 3: Partner strategies in alternating games with and without errors.** Partner strategies sustain cooperation in a Nash equilibrium. All such strategies are required to cooperate after mutual cooperation, such that the respective cooperation probability  $q_{CC}$  is equal to one. **a**, In the absence of errors, the remaining three cooperation probabilities can be chosen arbitrarily, subject to the constraints in Eq. (2). The resulting set of partner strategies takes the shape of a polyhedron. **b**, In the presence of errors, this polyhedron degenerates to a single line segment. This line segment comprises all strategies between Generous Tit-for-Tat (*GTFT*) and Stochastic Firm-but-Fair (*SFBF*). **c,d**, We compare these equilibrium results to evolutionary simulations. To this end, we record all strategies that emerge over the course of the simulation. Here, we plot the probability distribution of those strategies that yield at least 95% cooperation against themselves. Without errors, the probability distributions for  $q_{CD}$ ,  $q_{DC}$ ,  $q_{DD}$  are comparably flat. With errors, players tend to cooperate if they exploited their opponent in the previous round,  $q_{DC} \approx 1$ . Moreover, they cooperate with some intermediate probability after mutual defection,  $q_{DD} \approx 2/3$ . Both effects are in line with previous simulation studies<sup>22,23</sup>, and they confirm the theory. Simulations are run for  $b/c=3$ , and  $\varepsilon=0$  or  $\varepsilon=0.02$ . For the other parameter values and further details on the simulations, see **Methods**.



**Figure 4: Comparing evolution in the alternating and the simultaneous game.** To compare the two game versions, we have run additional evolutionary simulations. We systematically vary the benefit of cooperation, the population size, the selection strength, and the mutation rate. In addition, we vary how likely players make errors. Either they make no errors at all ( $\varepsilon = 0$ ), or they make errors at some intermediate rate ( $\varepsilon = 0.02$ ). **a**, In the absence of errors, there is virtually no difference between the simultaneous and the alternating game. Both games yield the same cooperation rates, and they respond to parameter changes in the same way. For the given baseline parameters, cooperation is favored for large benefits of cooperation, population sizes, and selection strengths. It is disfavored for intermediate and large mutation rates. **b**, With errors, the cooperation rates in the alternating game are systematically below the simultaneous game. The lower cooperation rates are related to our analytical result that no cooperative memory-1 strategy in the alternating game is evolutionarily stable. In contrast, in the simultaneous game with errors, *WSLS* can maintain cooperation<sup>36,49</sup>, it is evolutionarily stable<sup>35</sup>, and it readily evolves in evolutionary simulations (**Supplementary Fig. 1**). As baseline parameters we use a benefit of cooperation  $b=3$ , population size  $N=100$ , selection strength  $\beta=1$ , and the limit of rare mutations  $\mu \rightarrow 0$ <sup>61,62</sup>.

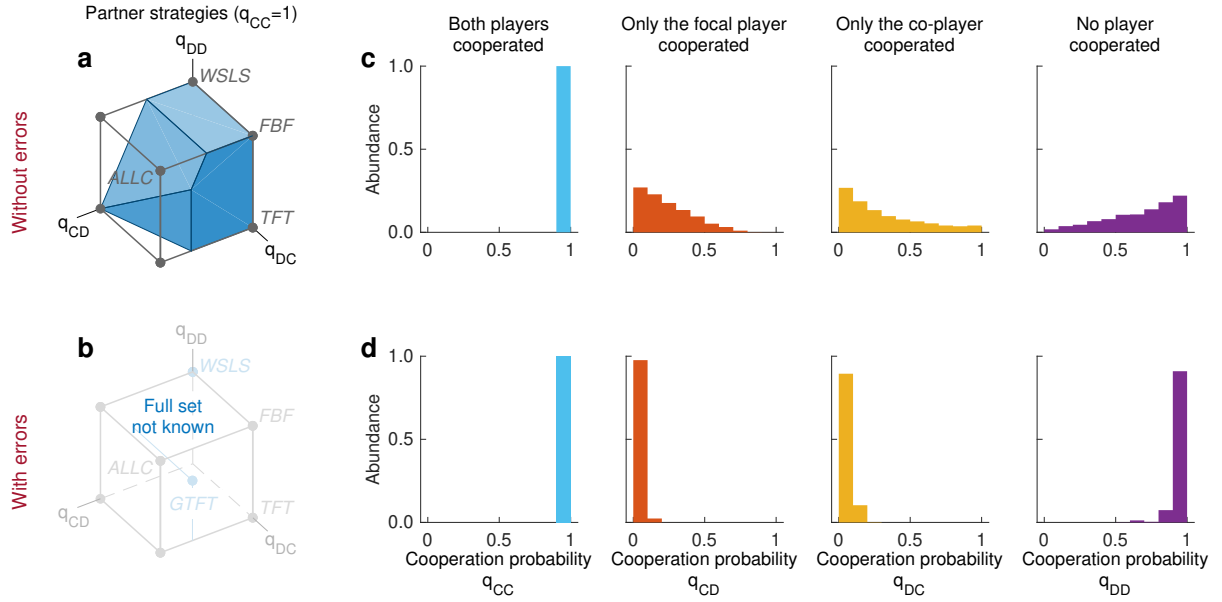




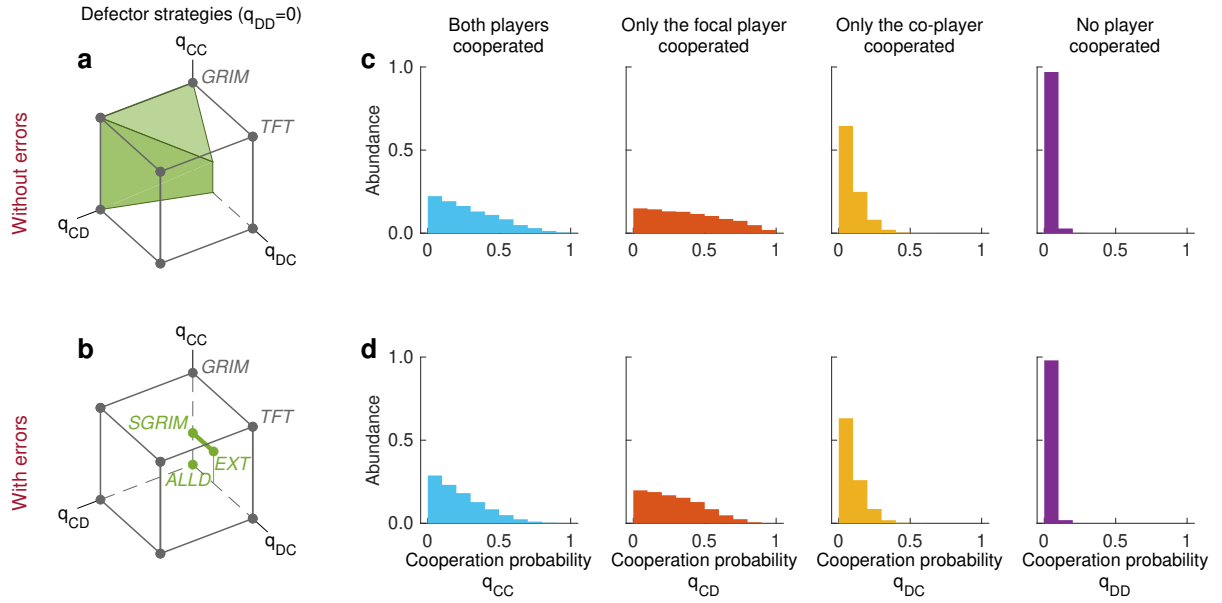
**Figure 5: Robustness of evolutionary results.** We have explored the robustness of our results with various model extensions. Here, we display results for three of them, illustrating the impact of finitely repeated games, of irregular alternating patterns, and of population structure. **a–c**, The baseline model assumes infinitely repeated games; here we show simulations for games with a finite expected length. If there are sufficiently many rounds, the simultaneous game again leads to more cooperation than the alternating game, and the evolving strategies are largely similar to the ones observed in the baseline model. **d–f**, The baseline model assumes that players move in a strictly alternating fashion. Instead, here we assume that after each player’s move, the other player moves with some switching probability  $s$ . The case  $s = 1$  corresponds to strict alternation, whereas  $s = 1/2$  represents a case in which the next player to move is completely random. We observe that irregular alternation patterns hardly affect which strategies players use to cooperate. However, it affects the robustness of these strategies. Overall, cooperation is most likely to evolve under strict alternation. **g–i**, Finally, instead of well-mixed populations, we consider games on a lattice. For the given parameter values, we observe that simultaneous games eventually lead to homogeneous cooperative populations. While this outcome is also possible for alternating games, some simulations also lead to the coexistence of cooperators and defectors (shown here in panel **h**). The evolving self-cooperative strategies are similar to the strategies that evolve in the baseline model. For a detailed description of these simulations, see **Methods** and **Supplementary Information**.

	Simultaneous game	Alternating game
Game without errors	<p><b>a</b></p> <ul style="list-style-type: none"> <li>The full set of partner strategies is <b>known</b>.</li> <li>Partner strategies exist for all <math>b &gt; c</math>.</li> <li>A strategy is a partner if and only if <math display="block">q_C = 1 \text{ and } q_D \leq 1 - \frac{c}{b}</math> </li> <li>None of these strategies is evolutionarily stable.</li> </ul>	<p><b>b</b></p> <ul style="list-style-type: none"> <li>We characterize partner strategies <b>here</b>.</li> <li>Partner strategies exist for all <math>b &gt; c</math>.</li> <li>A strategy is a partner if and only if <math display="block">q_C = 1 \text{ and } q_D \leq 1 - \frac{c}{b}</math> </li> <li>None of these strategies is evolutionarily stable.</li> </ul>
Game with errors	<p><b>c</b></p> <ul style="list-style-type: none"> <li>The full set of partner strategies is <b>known</b>.</li> <li>Partner strategies exist when <math>b &gt; c</math> and <math>\epsilon &lt; \frac{1}{2}\left(1 - \frac{c}{b}\right)</math>.</li> <li>The only partner strategy is <i>GTFT</i> with <math display="block">q_C = 1 \text{ and } q_D = 1 - \frac{c}{(1-2\epsilon)b}</math> </li> <li><i>GTFT</i> is not evolutionarily stable.</li> </ul>	<p><b>d</b></p> <ul style="list-style-type: none"> <li>We characterize partner strategies <b>here</b>.</li> <li>Partner strategies exist when <math>b &gt; c</math> and <math>\epsilon &lt; \frac{1}{2}\left(1 - \frac{c}{b}\right)</math>.</li> <li>The only partner strategy is <i>GTFT</i> with <math display="block">q_C = 1 \text{ and } q_D = 1 - \frac{c}{(1-2\epsilon)b}</math> </li> <li><i>GTFT</i> is not evolutionarily stable.</li> </ul>

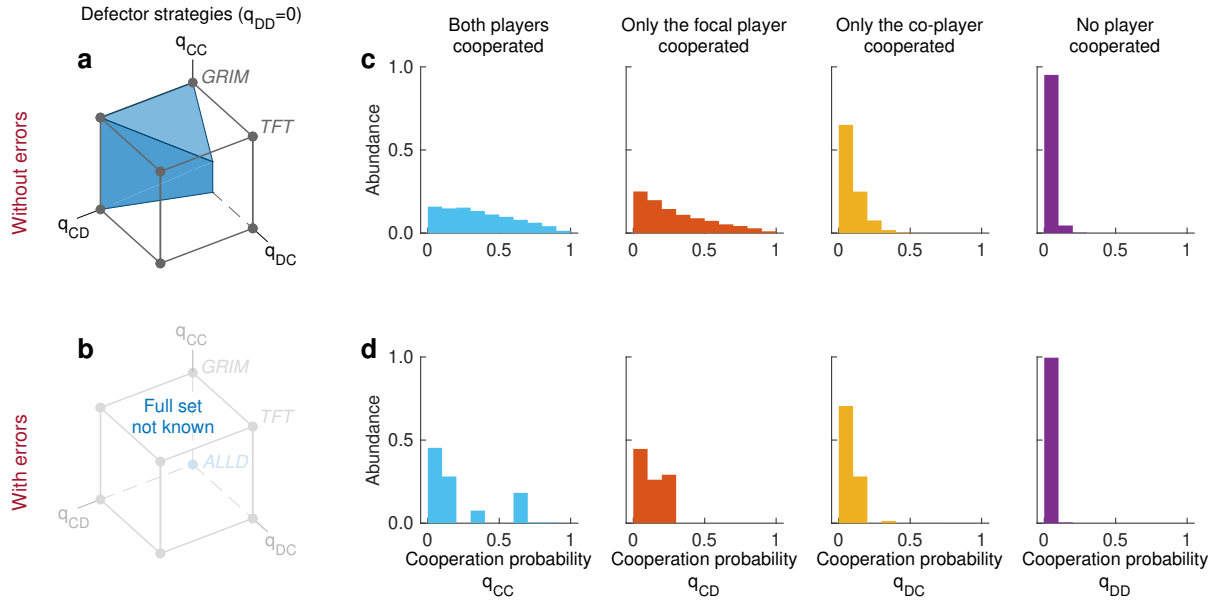
**Supplementary Table 1: A characterization of partners among the reactive strategies.** Here, we describe the set of partner strategies within the class of reactive strategies. Reactive strategies are a subset of memory-1 strategies. They consist of two conditional cooperation probabilities,  $q_C$  and  $q_D$ . The two probabilities describe how a player responds to a co-player's cooperation and defection, respectively. For reactive strategies, the alternating and the simultaneous game lead to the same payoffs<sup>22</sup>. As a result, also the partner strategies coincide in each case. With errors, *GTFT* is the only partner strategy. It can be neutrally invaded by *ALLC*, and hence it is not evolutionarily stable.



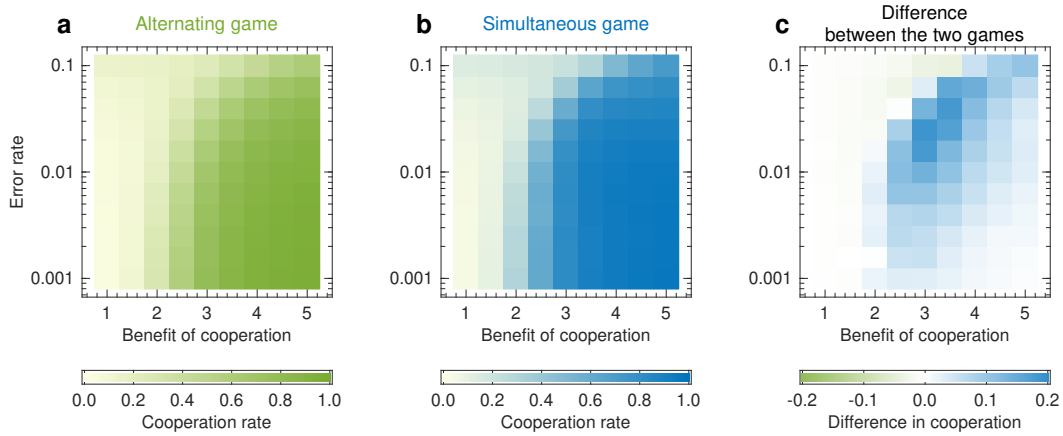
**Supplementary Figure 1: Partners in the simultaneous game.** Here, we present results analogous to those of **Fig. 3**, but illustrating the case of the simultaneous game. Without errors, the two models make the same predictions, and also the evolving cooperative strategies are similar. With errors, there are two differences. First, for the simultaneous game, there is no theory yet that characterizes all partner strategies. It is only known that particular cooperative strategies are stable under certain conditions. For example, *WSLS* is a partner strategy if  $b > 2c$ , provided the error rate is sufficiently small<sup>35</sup>. Another example of a partner strategy is *GTFT* (defined in the same way as in the alternating game; for reactive strategies, the two games are equivalent<sup>22</sup>). Second, our evolutionary simulations confirm that players predominantly maintain cooperation with *WSLS* in the simultaneous game with errors. Parameters are the same as in **Fig. 3**.



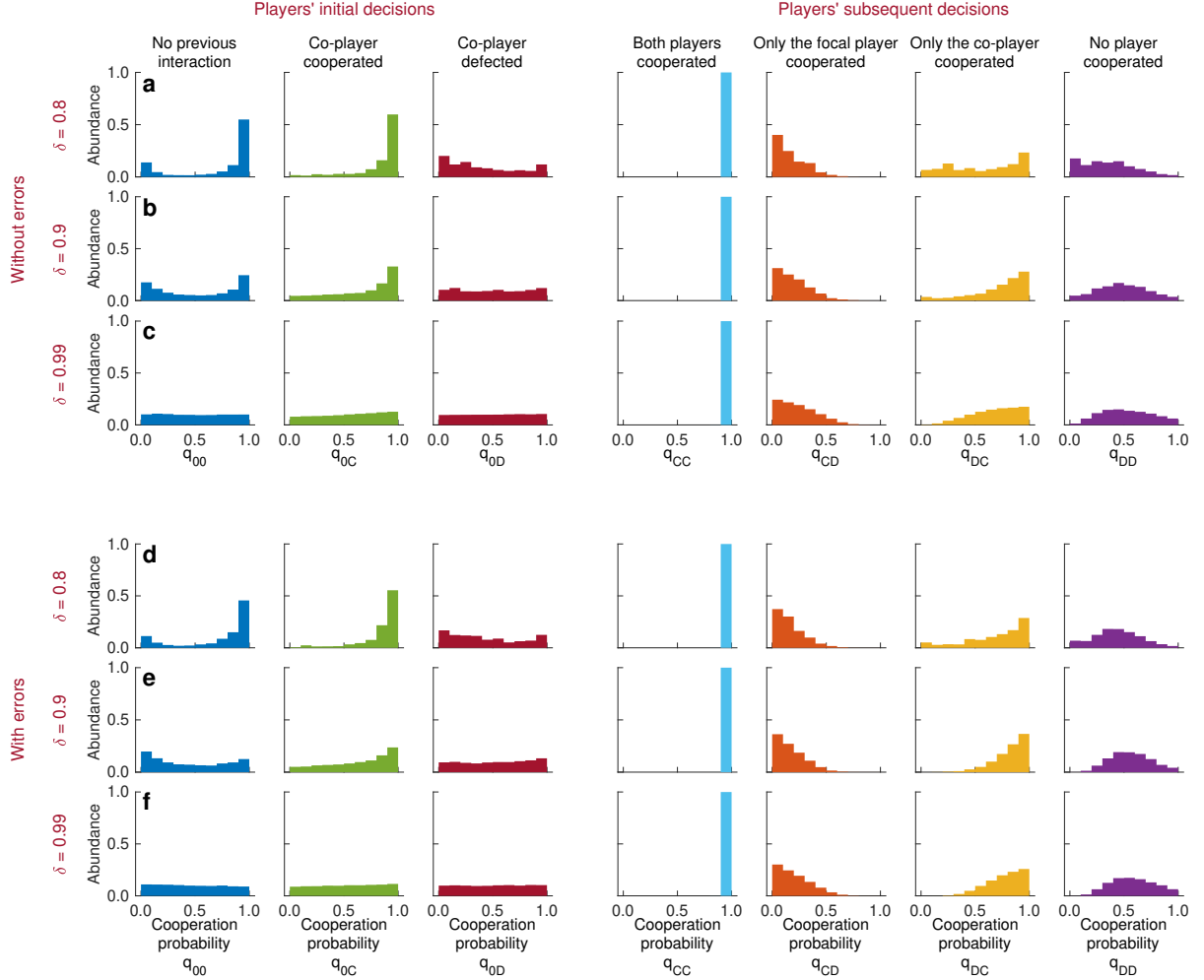
**Supplementary Figure 2: Defectors in the alternating prisoner's dilemma. a,b,** As we have done for partners, we have also characterized all defectors among the memory-1 strategies for the alternating game. We provide explicit conditions for the case without errors, see Eq. (3), and the case with errors, see **Supplementary Information**. With errors, there are two classes of defector strategies. First, there is the atomic class of unconditional defection (*ALLD*). Second, there is a line segment that connects a stochastic version of *GRIM* to the strategy *EXT*; the latter is a limiting version of the previously described extortionate strategies<sup>33,63–66</sup>. Parameters are the same as in **Fig. 3**.



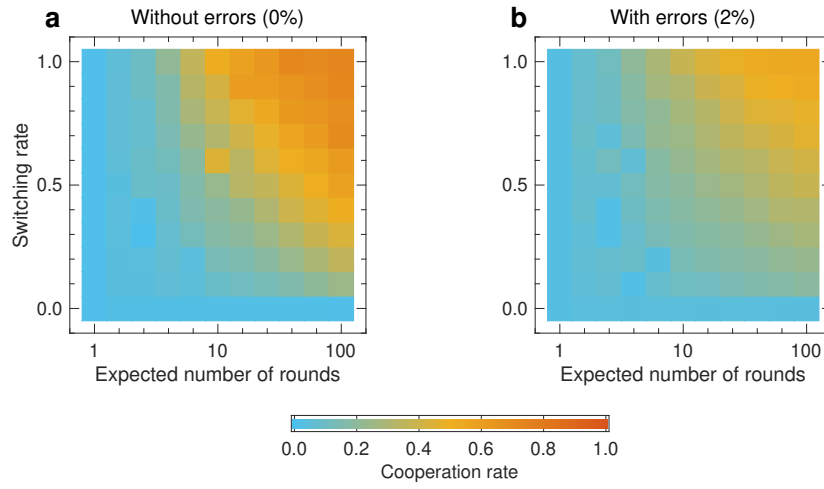
**Supplementary Figure 3: Defectors in the simultaneous game.** The figure is analogous to **Supplementary Fig. 2**, but for the case of the simultaneous game instead of the alternating game. For the simultaneous game with errors, there is no complete characterization of defector strategies as of yet. However, it is known that *ALLD* is a Nash equilibrium for all parameter values, because it simply reiterates the Nash equilibrium of the one-shot game. Parameters are the same as in **Fig. 3**.



**Supplementary Figure 4: Comparing the alternating and the simultaneous game across different error rates and benefit values.** We have run further simulations to explore the joint effect of the error rate and the benefit of cooperation, for both the alternating game (a) and the simultaneous game (b). In both cases, we observe high cooperation rates for high benefit values and sufficiently small error rates. c, Here, we plot the difference in cooperation rates between the simultaneous and the alternating game. This difference is small for small benefit values (where defection evolves in both settings). It is also small for large benefits when the error rate is small (for which nearly full cooperation evolves in both settings). In between, for intermediate benefit values and intermediate error rates, the simultaneous game yields systematically more cooperation than the alternating game. Baseline parameters are the same as in Fig. 4.

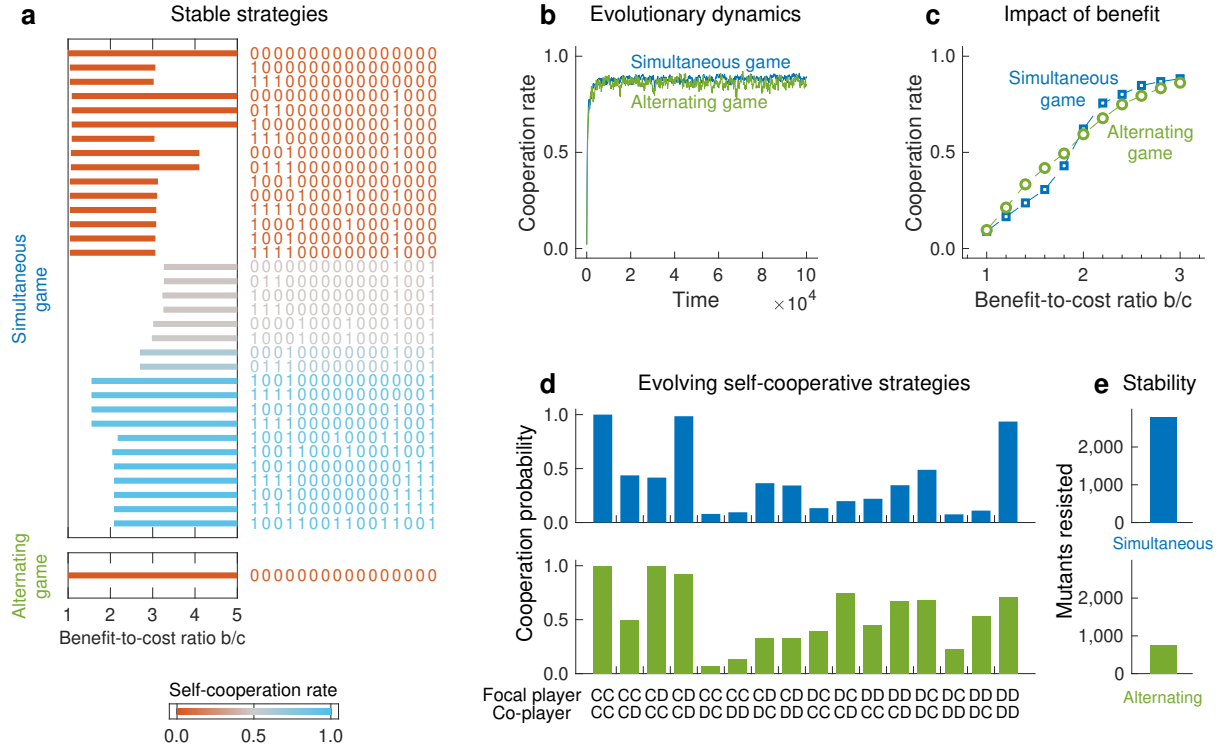


**Supplementary Figure 5: Self-cooperative strategies in the finitely repeated alternating game.** We explore which strategies the players use to maintain cooperation in finitely repeated alternating games. To this end, we consider two different error scenarios ( $\varepsilon = 0$  and  $\varepsilon = 0.02$ ), and three expected game lengths (5, 10, or 100 rounds, corresponding to a continuation probability of  $\delta = 0.8$ ,  $\delta = 0.9$ , and  $\delta = 0.99$ ). In each case, we run simulations and record those strategies that have a self-cooperation rate of at least 80%. Here, we show the distribution of these strategies. We observe the following regularities: (i) The players' first round behavior is only under selection when players interact for a few rounds. For  $\delta > 0.9$ , the distribution of the respective cooperation probabilities  $q_{00}$ ,  $q_{0C}$ ,  $q_{0D}$  is comparably flat. (ii) The conditional cooperation probabilities for all subsequent rounds generally resemble the evolving strategies of the baseline model. Apart from the parameters varied explicitly, parameters are the same as in Fig. 3.

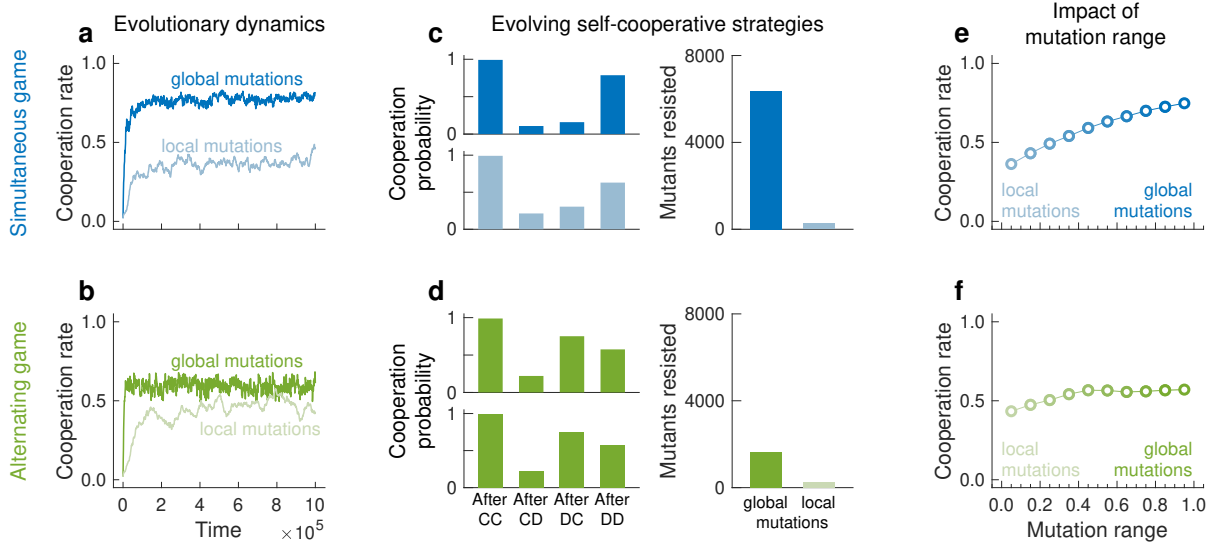


**Supplementary Figure 6: Cooperation in the finitely repeated game with irregular alternation patterns.** Using the same basic setup as in **Fig. 5**, we have explored how likely players are to cooperate in finitely repeated alternating games with irregular alternation patterns. To this end, we vary (i) the expected number of rounds, and (ii) the switching rate that measures how strictly players are to alternate. In addition, we consider two scenarios, depending on whether or not players commit implementation errors. As already indicated by **Fig. 5**, players are most likely to cooperate when there are no errors, when the number of rounds is large, and when players move in a strictly alternating fashion.





**Supplementary Figure 7: Stability and evolutionary dynamics of pure memory-2 strategies.** We have also run simulations for (infinitely repeated) simultaneous and alternating games when players have access to all 65,536 memory-2 strategies. These strategies are 16-dimensional vectors that take each player's last two actions into account (see **Supplementary Information** for details). **a**, In a first step, we have computed which of these strategies are evolutionarily stable for an error rate of  $\varepsilon = 0.02$ . Here, we display the respective strategies (encoded by the 16 integers on the right hand side), the range of  $b/c$  values for which these strategies are stable (indicated by the length of the lines), and the self-cooperation rates of these strategies (indicated by the color of the respective line). In the simultaneous game, there are many evolutionarily stable strategies, including strategies that yield almost full cooperation. In contrast, in the alternating game, *ALLD* is the only strategy that is evolutionarily stable for a positive range of  $b/c$  values. **b,c** Although only the simultaneous game allows for evolutionarily stable cooperation, simulations suggest that alternating games yield similar average cooperation rates. **d**, In a next step, we have recorded which strategies the players use to cooperate among themselves (for this simulation we again call a strategy self-cooperative if it yields a cooperation rate of at least 80% against itself). In the simultaneous game, the self-cooperative strategies resemble the previously reported all-or-none strategies<sup>35</sup>. Here, the two players are most likely to cooperate if they both cooperated in the last two rounds, if none of them cooperated in the last two rounds, or if they both cooperated in the last round but defected in the second-to-last round. In the alternating game, the players' conditional cooperation probabilities seem more irregular. **e**, We have also computed how robust self-cooperative strategies are, by recording how many mutant strategies it takes on average to successfully invade into a resident population of self-cooperators. As expected from our evolutionary stability analysis, self-cooperative strategies are more robust in simultaneous games. For details, see **Methods** and **Supplementary Information**.



**Supplementary Figure 8: Evolutionary dynamics under local mutations.** The previous simulations assume that mutations are global: mutant strategies can be arbitrarily far away from the resident strategy. Here we compare this scenario with the case of local mutations, where mutant strategies are required to be in a small neighborhood of the resident strategy. We measure the size of this neighborhood by the mutation range  $m$ . The mutation range reflects by how much the mutant's conditional cooperation probabilities are allowed to differ from the resident strategy. Unless noted otherwise, we use  $m=0.05$ . **a,b**, In both the simultaneous and the alternating game, local mutations lead to less cooperation. However, the effect is more notable in the simultaneous game. **c,d**, Local mutations do not affect which strategies the players use on average to maintain cooperation. However, they affect how robust these strategies are. Under local mutations, all mutants have approximately the same fitness as the resident. As a result, the evolutionary competition is almost neutral; on average, it thus takes an order of  $N$  mutants to invade any given resident population (here, the population size is  $N=100$ ). **e,f**, We have repeated these simulations for different mutation ranges. The mutation range has a strong effect on cooperation in the simultaneous game (where evolutionarily stable cooperation is possible). It has a comparably weak effect in the alternating game (in which no evolutionarily stable strategy exists that leads to full cooperation).