Received: 26 May 2023 | Revised: 7 September 2023 | Accepted: 13 September 2023
DOI: 10.1111/eth.13407

REVIEW

ethology WILEY

Direct reciprocity among humans

Charlotte S. L. Rossetti | **Christian Hilbe**

Max Planck Research Group on the Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, Plön, Germany

Correspondence
Charlotte S. L. Rossetti, Max Planck Research Group on the Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, Plön 24306, Germany.
Email: rossetti@evolbio.mpg.de

Funding information
H2020 European Research Council, Grant/Award Number: 850529 (E-DIRECT)

Editor: Manon Schweinfurth

Abstract

Direct reciprocity is the tendency to repay others' cooperation. This tendency can be crucial to maintain cooperation in evolving populations. Once direct reciprocity evolves, individuals have a long-run interest to cooperate, even if it is costly in the short run. The major theoretical framework to describe reciprocal behavior is the repeated prisoner's dilemma. Over the past decades, this game has been the major workhorse to predict when reciprocal cooperation ought to evolve, and which strategies individuals are supposed to adopt. Herein, we compare these predictions with the empirical evidence from experiments with human subjects. From a theory-driven perspective, humans represent an ideal test case, because they give researchers the most flexibility to tailor the experimental design to the assumptions of a model. Overall, we find that theoretical models describe well in which situations people cooperate. However, in the important case of "indefinitely repeated games," they have difficulties to predict which strategies people use.

KEYWORDS
direct reciprocity, evolution of cooperation, evolutionary game theory, human behavior, reciprocal altruism

1 | INTRODUCTION

Cooperation is a fundamental component of many social interactions (Raihani, 2021; Rossetti et al., 2022). It occurs when individuals share food and other commodities (Schweinfurth & Taborsky, 2018), contribute to a collective action (Olson, 1971), or when they use public resources responsibly (Hardin, 1968; Ostrom, 1990). The common pattern behind these examples is that individuals incur a personal cost to benefit others. Such seemingly altruistic behaviors warrant an explanation: if cooperation is to evolve, it needs to give a fitness advantage to the cooperating individual or its kin (Colman, 2006). Importantly, however, such a fitness advantage does not need to arise immediately. Instead it suffices if there is some advantage eventually, over the course of an individual's lifetime. This insight provides the basis for direct reciprocity (Trivers, 1971), one of the key mechanisms for cooperation (Nowak, 2006; Sigmund, 2010). When individuals interact in stable groups, their cooperative acts today may lead other group members to cooperate with them in future. Once future benefits are sufficiently valuable, (conditional) cooperation is what evolution selects for.

In nature, cooperation can come in various degrees, and it can involve many individuals. Yet when modeling behavior, it is often useful to consider idealized scenarios that capture a behavior's central features in the simplest possible way. One frequently used paradigm to study cooperation is the prisoner's dilemma (Rapoport & Chammah, 1965). It describes an interaction among two individuals (in theoretical studies, the prisoner's dilemma is often described as a "game," individuals are referred to as "players," and outcomes are called "payoffs"; however, the framework covers scenarios that are less innocent than these names might suggest). The rules of the interaction are as follows. Each individual can choose to cooperate or to defect. Mutual cooperation yields the highest payoff to the pair, yet defection yields a higher payoff to each individual. Because choices are made independently, the only reasonable and consistent outcome of the prisoner's dilemma—the only Nash equilibrium—is mutual defection. The prisoner's dilemma is widely used because it captures the essence of cooperation: the conflict between self-interest and group interest. At the same time, it is arguably the most simple model to do so: there are only two players (instead of many), and players can only choose among two discrete actions (there are no different shades of good and bad behavior).

While defection is the only equilibrium in a single interaction, predictions change when players interact over multiple rounds. In that case, players can adopt reciprocal strategies to enforce cooperation. They can cooperate with other cooperators, and they can stop cooperating against defectors. Importantly, this form of reciprocity can evolve even when individuals do not consciously compute the payoffs of their actions. In the end, evolution leads them to behave *as if* they did, simply because it wipes out behaviors that are not well-adapted to an individual's environment. After decades of research, there is by now a vast theoretical literature on evolution in the repeated prisoner's dilemma (Glynatsi & Knight, 2021). This literature explores how evolving cooperation rates depend on the parameters of the game and on the exact setup of the evolutionary process (Hilbe et al., 2018).

Importantly, these models in evolutionary game theory typically take an ultimate, not a proximate, perspective. They ask in which kinds of environments cooperative behavior would be adaptive. To this end, the models neglect any specific emotions that individuals might feel when making their decisions, or any values that individuals might hold. Rather the models ask which kinds of strategies allow for stable cooperation in a given environment, irrespective of the proximate mechanisms that might lead individuals to implement those strategies. Real behavior might not perfectly resemble the strategies predicted by this theory. Yet, we would hope that evolutionary theory can give us some clues on which behavioral patterns are essential for reciprocity to succeed. At the same time, it should be noted that evolutionary models do not have the aim of exactly predicting cooperation levels. Rather they allow us to explore which qualitative features of an interaction are favorable to cooperation and which are not.

In this article, we compare the theoretical predictions for the prisoner's dilemma with empirical evidence from behavioral experiments with humans. Humans represent an ideal test case for evolutionary models for various reasons. First, humans develop the capacity for reciprocity already at an early age (Wörle & Paulus, 2019), and a majority of adults engage in behaviors consistent with conditional cooperation (Fischbacher et al., 2001; Grujic et al., 2014). Second, online and laboratory experiments with human subjects are straightforward to implement and comparably cheap. Third, the experimental design and the instructions can be easily tailored to explore the impact of different payoff parameters, stopping conditions, and learning horizons. For some results, it is also useful that humans are capable to respond to hypothetical scenarios. For example, by letting participants interact with computerized opponents, one can explore how they would react to certain predefined strategies that are relevant for the theoretical literature (Crandall et al., 2018; Hilbe et al., 2014; Xu et al., 2016). Of course, the resulting insights on human subjects cannot be easily extrapolated to other species. Nor can they be easily extrapolated to human interactions in real life, where reciprocity is more difficult to quantify (for a recent exception, see Frank et al., 2018). However, due to the flexibility of experiments with human subjects, they can serve as a first test case to determine which models of reciprocity might be sensible *in principle*. We use these insights to reflect on the success of evolutionary models, and to identify open problems that require more work.

The remainder of this article is organized as follows. In the next section, we briefly review the theoretical literature on the repeated prisoner's dilemma. We then comment on typical experimental implementations and describe their impact on observed average cooperation rates. Afterwards, we review common conditional strategies observed in human reciprocal interactions, and we discuss cognitive constraints and their impact on reciprocity. Finally, we provide a brief overview of reciprocal interactions captured by models different from the standard prisoner's dilemma.

2 | THEORETICAL BACKGROUND

2.1 | The repeated prisoner's dilemma

The prisoner's dilemma is a game among two players who independently decide whether to cooperate (C) or to defect (D), as illustrated in Figure 1a. Mutual cooperation yields a *reward* of R to both players, whereas mutual defection results in the *punishment* payoff P. If one player defects whereas the other cooperates, the defector obtains the *temptation* payoff T whereas the cooperator ends up with the *sucker's payoff* S. For the game to be a prisoner's dilemma, the payoffs have to satisfy the inequalities $T>R>P>S$. When these inequalities hold, game theory tells us that the rational choice for both players is to defect although mutual defection yields a lower payoff than mutual cooperation. This is because defection is the only "safe" choice where both players cannot do anything else that will make them better off. In addition to the above inequalities, most models also assume that $2R>T+S$. This latter assumption ensures that it is the symmetric outcome of mutual cooperation that yields the highest total payoff, rather than the asymmetric outcome in which one player cooperates and the other defects.

There are two particular instantiations of the prisoner's dilemma that are often used as baseline examples. One is based on the payoffs $R=3$, $S=0$, $T=5$ and $P=1$ (Figure 1b). From a theoretical viewpoint, there is nothing special about these particular parameter values, other than that they were used in the seminal study of Axelrod and Hamilton (1981). From an experimental viewpoint, however, it must be noted that all these payoffs are non-negative. While mathematical

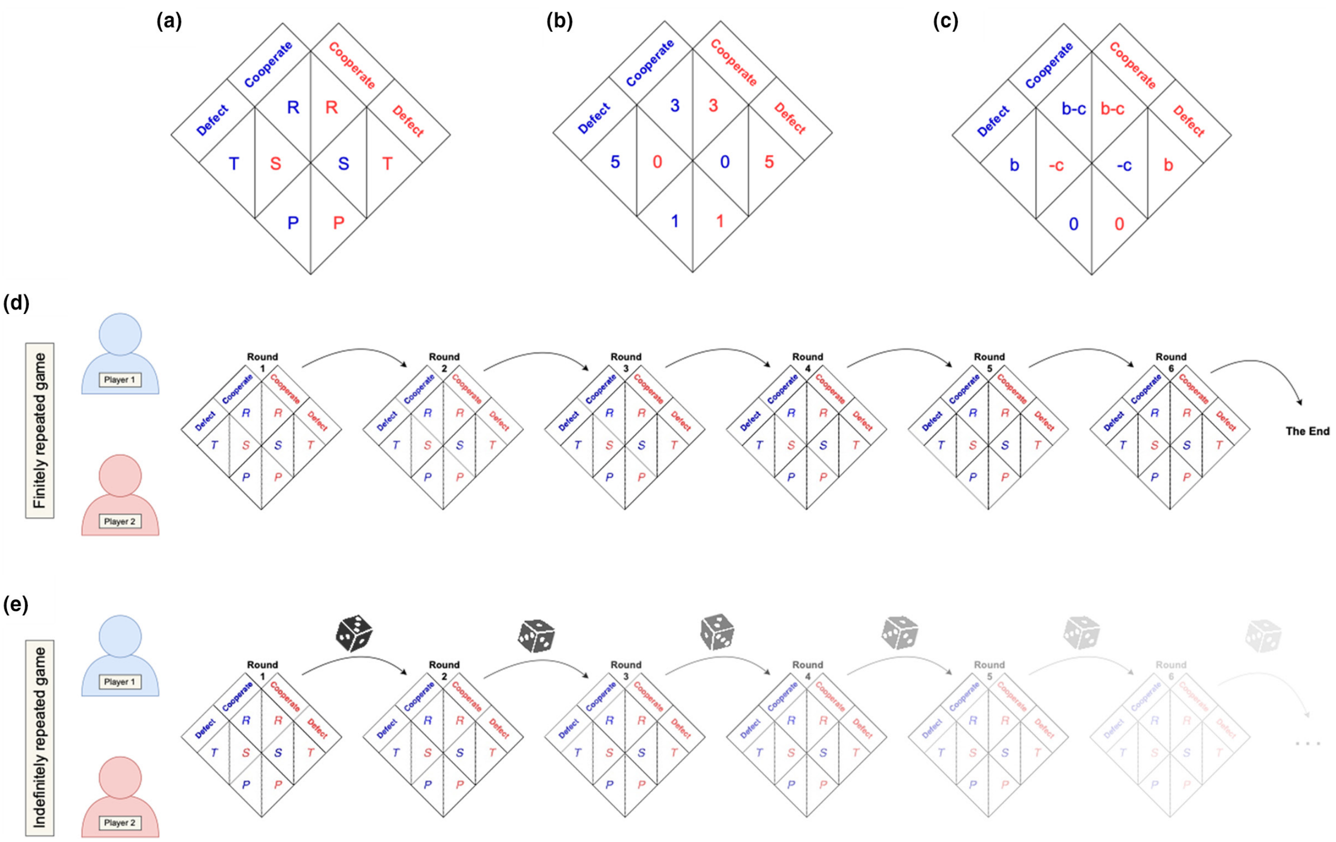

FIGURE 1 Basic setup of the repeated prisoner's dilemma. (a) In the prisoner's dilemma, two individuals (here depicted as blue and red) independently decide whether to cooperate or defect. Mutual cooperation gives a reward R to both, whereas mutual defection yields the lower punishment payoff P to both. If one player cooperates and the other defects, the defector gets the highest payoff T (temptation), whereas the cooperator gets the smallest payoff S (the sucker's payoff). (b) Axelrod and Hamilton (1981) studied a particular variant of this game with payoffs $T=5$, $R=3$, $P=1$, and $S=0$, which has become a baseline since. (c) Another popular representation of the prisoner's dilemma is the donation game, in which payoffs are framed in terms of benefit b and cost c of cooperation. In the theoretical literature, it is common to distinguish two variants of repeated games: (d) In the finitely repeated prisoner's dilemma, the two players interact for a known number n of rounds. In particular, in the last round, players are aware that no further interactions will occur. (e) In the indefinitely repeated prisoner's dilemma, there is a constant chance that a further round occurs. In particular, players can never be sure that they will not interact again.

predictions typically only depend on the relative magnitudes of payoffs, not on their absolute values or signs, humans are known to be sensitive to negative framing. The other instantiation is the so-called donation game with payoffs $R=b-c$, $S=-c$, $T=b$, and $P=0$, where $b>c>0$ denote the benefit and the cost of cooperation, respectively (Figure 1c). While these two instantiations satisfy all of the above inequalities, they do not generate the entire space of all prisoner's dilemmas (which instead would require using the general payoffs R, S, T, P). However, in many cases the specific payoffs of Axelrod and of the donation game are easier to work with, which explains their wide use in many evolutionary models (Hilbe et al., 2018).

To explain direct reciprocity, we are interested to see what happens when the game is repeated. In a repeated prisoner's dilemma, we now have a social interaction with multiple encounters, such that players interact for several rounds (such iterated interactions are sometimes referred to as “supergames”). From a theoretical perspective, it is useful to distinguish two different kinds of repeated interactions. They are referred to as the finitely and the indefinitely repeated game, respectively. In the finitely repeated game, the two players interact for a commonly known number n of rounds (Figure 1d). Perhaps somewhat surprisingly, the standard prediction for finitely repeated game is the same as for the (one-shot) prisoner's dilemma (the one with $n=1$). This result follows from *backward induction*: in the very last round n, players no longer have any incentive to cooperate, and hence they should both defect. However, given both players defect in round n anyway, it becomes optimal to already defect in round $n-1$, and by the same logic, in all previous rounds. This race-to-the-bottom logic no longer applies in the indefinitely repeated game (Figure 1e). Here, there is no commonly known last round. Instead, after any interaction, there is always a probability $\delta>0$ of a further encounter. According to an equivalent interpretation, one may also imagine two players who interact for infinitely many rounds, but who discount future payoffs with a discount factor of δ. For this reason, indefinitely repeated games are sometimes also referred to as “infinitely repeated games” (Dal Bó and Fréchette, 2018), even if $\delta<1$. Once there is no predetermined last round, reciprocal cooperation becomes feasible. Hence it is the indefinitely repeated game that is considered in most (but not all) theoretical studies on the evolution of reciprocity.

When we study behavior in games, we look at what strategies players use. In a one-shot game, there are as many pure strategies as there are actions: players can either cooperate or defect. In contrast, when the game is repeated, the number of strategies can be vast (it becomes infinite when the game is itself infinite). This is because strategies for the repeated game correspond to contingent plans of action. They tell the player what to do in any round, depending on what happened in all previous rounds. For example, always defect (*ALLD*) is a strategy. Choose at random is also one. Cooperate all the time but defect every fourth round as well. Only some strategies are interesting, either because they are played by human subjects or because of their theoretical properties. In particular, researchers tend to look at conditional strategies. Unlike the examples given just above, this set of strategies take into account the co-player's previous behavior. For example, a player may cooperate as long as the other co-player does, then defect every time. This strategy is called *GRIM* (Sigmund, 2010). Another example is the strategy Tit-for-Tat (*TFT*), where players simply copy what the other player did on the previous round.

Because the space of possible strategies of the prisoner's dilemma is enormous, it is common in the evolutionary literature to artificially restrict the space of strategies that players can use. For example, some studies assume that players only react to the outcome of the last round, or more generally the last k rounds (Hauert & Schuster, 1997; Hilbe et al., 2017; Martinez-Vaquero et al., 2012; Murase & Baek, 2020; Nowak & Sigmund, 1993). Some other studies assume that individual strategies need to be representable by a so-called finite-state automaton (van Veelen et al., 2012). The states of such an automaton can be thought of as the players' different mental states (such as being "satisfied" or being "angry"). The players' states in the current round determine which actions they choose, which in turn determine the players' state in the next round. A few instances of such strategies, including the strategies *ALLD*, *GRIM*, and *TFT*, are described in Table 1. Restricting the players' feasible strategies (to either have finite memory or finitely many states) serves two purposes. On the one hand, it captures that humans rarely act as perfectly calculating machines that condition their behavior on the entire previous history of interactions. On the other hand, these restrictions allow researchers to more efficiently explore which strategies might evolve. For example, only when players are restricted to choose from a reasonably small set, one can hope to explore the dynamics with computer simulations.

A final modeling assumption that is often made is that people may commit errors. For example, they may commit implementation errors: in situations in which players would usually cooperate, they might instead defect with some probability ε, possibly because of a "trembling hand" (Selten, 1975). Alternatively, it is sometimes assumed that individuals misremember past events, possibly due to a "fuzzy mind" (Stevens et al., 2011). Again, the assumption of errors

TABLE 1 Some archetypical strategies of the repeated prisoner's dilemma.

Name	Desciption	Type	Memory length	Reference
Always Defect (AllD)	Play defect in all rounds	Unconditional	0	
Always Cooperate (AllC)	Play cooperate in all rounds	Unconditional	0	
Tit-for-Tat (TFT)	Cooperate on the first round, then copy the previous action of the other player	Conditional	1	Rapoport and Chammah (1965), Axelrod and Hamilton (1981)
Grim trigger (Grim)	Cooperate on the first round and continue to cooperate as long as the other player cooperates. If the other player defects once, then defect forever.	Conditional	1	
Win-Stay Lose-Shift (WSLS)	Cooperate on the first round. Then if both players chose the same action in the previous round, choose cooperate. If both player chose different actions in the previous round, choose defect	Conditional	1	Kraines and Kraines (1993), Nowak and Sigmund (1993)
Suspicious Tit-for-Tat (STFT)	Defect on the first round, then copy the previous action of the other player	Conditional	1	Boyd and Lorberbaum (1987)
Tit-for-Two-Tats (TF2T)	Cooperate on the first round. Then play TFT but only defect after two defections	Conditional	2	
Cumulative Reciprocity (CURE)	Cooperate on the first round. Then on each round, add one point to your tally if the other player defected when you cooperated. Remove one point if you defected when the other player cooperated. Neither remove nor add points if both players did the same. When the tally exceeds some predetermined threshold, say 3, defect	Conditional	∞	Li et al. (2022)

Note: We describe some strategies that have been highlighted by the previous theoretical literature, either because they have important properties, or because of they represent interesting extreme cases.

serves two purposes. On the one hand, it makes models more realistic. After all, chance events do sometimes interfere with people's decisions: sometimes we misinterpret an action, or we simply forget what we meant to do, or we meant to do that in another interaction. As a response, people seem to have developed ways to cope with these different kinds of noise (Tazelaar et al., 2004; Van Lange et al., 2002). On the other hand, models with errors are sometimes easier to tract mathematically and statistically (Sigmund, 2010). For example, without noise it can be difficult to infer a player's strategy solely based on the player's previous sequence of actions. This problem arises, for example, when two players both cooperate in all rounds. Such a sequence of actions is consistent with the assumption that both players are unconditional cooperators (*ALLC*). However, it is equally consistent with the assumption that both players cooperate conditionally (e.g., *TFT* or *GRIM*). Games with noise make it easier to distinguish these two cases: once one player defects (possibly by mistake), the other player can show her true colors.

2.2 | Theoretical predictions

After defining the rules and parameters of the game, we briefly discuss what kind of predictions have been derived from this model. These predictions can be based on several different ways how to analyze the repeated prisoner's dilemma, see Box 1. In the following, we summarize the general patterns that follow from this analysis, both for the finitely and for the indefinitely repeated game. In each case, we ask: how does a given parameter or assumption affect cooperation? In addition, we ask: which strategies are players predicted to adopt?

For the finitely repeated game, we have noted earlier that backward induction predicts that players fully defect eventually. This equilibrium prediction holds for all parameters (i.e., it is independent of the exact payoffs, or the exact number of rounds). There are, however, alternative models that predict some cooperation to emerge. These alternative models are based on the assumption that there is always a positive chance that a given co-player is conditionally cooperative—either because the co-player has social preferences (Kreps et al., 1982), or because such strategies are occasionally introduced by mutations (McNamara et al., 2004). Once there is even a small chance that the opponent might cooperate, conditional cooperation can become self-enforcing: Rather than trying to pre-empt the co-player's defection, it becomes rational to adopt a conditionally cooperative strategy, and to only defect once the co-player did so. This mechanism can lead to a substantial increase in predicted cooperation rates, especially if cooperation yields high benefits and if players interact for many rounds.

For the indefinitely repeated game, evolutionary and equilibrium arguments suggest that game parameters should affect cooperation in intuitive ways, see also Table 2. For example, the larger the benefit-to-cost ratio b/c, the more profitable cooperation becomes, and hence individuals should be more likely to cooperate (Akin, 2016; Stewart & Plotkin, 2015). A similar argument holds for

BOX 1 Theoretical methods to explore optimal play in the repeated prisoner's dilemma. Most previous research uses one of three different methods to explore optimal behavior in the repeated prisoner's dilemma: equilibrium analysis, computer tournaments, or evolutionary simulations.

Equilibrium analysis is a direct application of game theory and uses analytical methods to characterize which Nash equilibria are possible (Mailath & Samuelson, 2006). These equilibria are important because they give us some indication about which outcomes may occur in principle (strategies that are not equilibria are unlikely to persist). In the case of indefinitely repeated games, however, the equilibrium approach is surprisingly inconclusive. The celebrated "folk theorem" guarantees that almost any outcome might arise as an equilibrium if only the continuation probability is sufficiently large. The only requirement is that each player at least receives the mutual defection payoff P (see, e.g., Fudenberg & Maskin, 1986). In some cases, however, the required continuation probability might be prohibitively large in practice.

Round-robin tournaments represent another way to gain insights into the repeated prisoner's dilemma. Here, the assumption is that we can pit all strategies against each other and see which ones finish with the highest payoffs. This approach has been pioneered by Axelrod and Hamilton (1981), who found Tit-for-Tat (*TFT*) to succeed. Their study has since been repeated (and challenged) by several other groups (Knight et al., 2016; Rapoport et al., 2015). In particular, whether or not *TFT* succeeds depends on the strategies that are allowed to take part in the tournament, and on the game's parameters—such as the error rate (Glynatsi & Knight, 2020).

Finally, through *evolutionary simulations*, researchers can test which strategies emerge in evolving populations. Here, researchers assume that individuals repeatedly play against other population members, and successful players are more likely to reproduce. By exploring which strategies evolve eventually, researchers aim to identify behaviors that optimally support cooperation. Such evolutionary simulations often predict that *WSLS* or related strategies succeed (Hauert & Schuster, 1997; Martinez-Vaquero et al., 2012; Nowak & Sigmund, 1993; Szolnoki & Perc, 2014). The evolutionary approach is naturally connected to the other two. For example, in large populations with strong selection and rare mutations, the strategies that emerge correspond to the Nash equilbria of the game (Stewart & Plotkin, 2014). On the other hand, when mutations are frequent, such that all strategies are played in almost equal frequencies, evolution favors the strategy that would also succeed in the round-robin tournament (Tkadlec et al., 2023).

TABLE 2 Basic predictions for the repeated prisoner's dilemma.

Parameters	Effect	Description	Theoretical findings	Experimental evidence
Benefit	As the benefit of cooperation increases, cooperation increases	For a given continuation probability, the game's payoffs determine whether or not cooperation is feasible. Intuitively, both theory and experiments find that the higher the benefit in comparison to the cost, the more cooperation	van Veelen et al. (2012), Akin (2016), Stewart and Plotkin (2015)	Dal Bó and Fréchette (2019), Gill and Rosokha (2020)
Continuation probability	As the probability of another round increases, cooperation increases	Reciprocity is strongly influenced by the length of the interaction, or the likelihood that the current interaction will continue. High continuation probabilities lead to higher cooperation rates, especially for payoffs in equilibrium	Hilbe et al. (2015), Schmid et al. (2021)	Dal Bó and Fréchette (2018)
Error rate	As noise increases, cooperation first increases, then it decreases	Several studies already find a decrease in cooperation at small levels of noise. However, when players are allowed a longer memory (take more than just the last round into account), they learn to be more forgiving to compensate this issue, maintaining cooperation	Zhang (2018)	Aoyagi et al. (2003), Fudenberg et al., 2012, Li et al. (2022)

Note: In evolutionary models of the repeated prisoner's dilemma, the main parameters are the payoffs, the continuation probability, and the error rate. Here we summarize their predicted effects. The two last columns provide some references for the suggested relationships.

the continuation probability δ. The more likely it is that people interact for many rounds, the more important it becomes to maintain cooperative relationships, and hence cooperation should increase (e.g. Hilbe et al., 2015). The effect of errors is predicted to be ambivalent. Small error rates ε can sometimes enhance cooperation (Zhang, 2018), but frequent errors rates tend to be detrimental (Schmid et al., 2021). Moreover, cooperation can be further promoted if interactions are assorted rather than well-mixed (that is, when players are more likely to encounter co-players with the same strategy, van Veelen et al., 2012).

On the level of predicted strategies, there is a curious mismatch in predictions. Studies based on round-robin tournaments (when each contestant, here strategy, competes one-to-one with all others) often predict reciprocal strategies like *TFT* to be most successful (e.g., Axelrod & Hamilton, 1981). *TFT* cooperates if and only if the co-player did so in the previous round. This strict form of reciprocation can be advantageous in heterogeneous populations; by using *TFT*, a player can enforce that outcomes are fair, no matter what strategy the opponent adopts (Press & Dyson, 2012). On the other hand, *TFT* is very sensitive to errors. When two *TFT* players interact, already one (mistaken) defection is sufficient for mutual cooperation to break down. For this reason, studies based on evolutionary simulations often find that *TFT* only plays a transient role, and that players eventually learn to adopt a strategy of win-stay lose-shift (*WSLS*, Nowak & Sigmund, 1993). *WSLS* prescribes to repeat the previous action if the player's payoff was at least R, and to switch to the opposite action otherwise. Compared to strict reciprocation, *WSLS* has the strong advantage that it is robust with respect to errors. Indeed, even when one player defects by mistake, two *WSLS* players recover mutual cooperation after two rounds. Due to this property the strategy of *WSLS* is a Nash equilibrium in games with errors, whereas *TFT* is not (Hilbe et al., 2017). As a result, most evolutionary simulations predict that individuals should use *WSLS*, not *TFT*, to enforce cooperation.

3 | IMPACT OF DESIGN CHOICES AND PARAMETERS ON HUMAN COOPERATION

After discussing the central predictions of the theoretical literature, we compare them to the experimental evidence. Herein, we focus on data from controlled experiments with human subjects. These subjects have either been invited to interact in games in a laboratory, or they have been recruited through online platforms like Amazon Turk or Prolific (Horton et al., 2011). In each case, participants are asked to repeatedly make decisions in a repeated prisoner's dilemma and they are paid in proportion to their performance in the game (for similar evidence on the repeated public goods game, see for example Fischbacher & Gächter, 2010). Moreover, in some of the studies, individuals do not only engage in one repeated prisoner's dilemma (one supergame). Rather they consecutively act in several supergames with changing partners. In this way, the corresponding studies can disentangle two concurrent effects that both lead to behavioral change: strategic conditional play (within each supergame) and learning (across supergames).

3.1 | Dynamics in the finitely repeated prisoner's dilemma

For our summary of the experimental literature, we start with the finitely repeated prisoner's dilemma. Here, participants know the length of the game beforehand. Experimental outcomes seem to critically depend on how many rounds participants play. When there are only a few rounds, like two rounds or four, people seem to learn the logic of backward induction. As a result, they eventually start to defect early on in the game (Dal Bó, 2005). This picture changes, however, once there are more rounds. In that case, subjects seem to robustly cooperate until 3–4 rounds prior to the known ending of the game; only then there is a notable drop in cooperation (Cooper et al., 1996; Embrey et al., 2018). Cooperation does not seem to further unravel even if subjects have many opportunities to learn the specifics of the game. In an online game, participants interacted in 10-round prisoner's dilemmas with changing co-players over 20 days. Even by the end of the experiment, cooperation rates in each game remained high until round eight (Mao et al., 2017). Such results impose strong limits on backward induction. However, they agree with models suggesting that people cooperate conditionally because they assume others might do so as well (Kreps et al., 1982; McNamara et al., 2004). In line with this view, cooperation is even more pronounced if participants are told that some of their interactions may take place against computerized opponents who implement TFT (Andreoni & Miller, 1993).

3.2 | Dynamics in the indefinitely repeated prisoner's dilemma

Next, we consider experiments on the indefinitely repeated prisoner's dilemma. After every round, the game will stop there with some (known) probability $1-\delta$, or continue for at least one more round with probability δ. If the game continues, the same termination rule applies to the new round. A probability of $\delta=0$ means there is no other round, whereas a probability of 1 means there will be another round for sure. A probability of 0.5 means that there is a 1/2 chance of another round, and hence the expected number of rounds is $1/(1-\delta)=2$. All existing theoretical models suggest that cooperation ought to become more likely as δ becomes larger. Confirming this basic expectation, Dal-Bó and Fréchette (2018) find in an analysis of 15 new studies that cooperation increases with the probability of another round.

In addition to the continuation probability, cooperation is predicted to depend on the exact payoff parameters of the prisoner's dilemma. Most often payoffs are chosen to test a mathematical model (Roth & Murnighan, 1978). A specific combination of continuation rule and payoff matrices allow for different possible equilibria. Economists are interested to see how the availability of different equilibria affects behavior (Dal-Bó & Fréchette, 2018). But even if the set of possible equilibria is unchanged, different payoff matrices can lead to different choices from players. For example, research suggests that cooperation is more abundant when it is risk-dominant; in this case, risk-dominance means that players prefer to cooperate when they think it is equally likely that the co-player adopts *ALLD* or *GRIM* (Blonski et al., 2011; Dal Bó & Fréchette, 2011). In particular, a higher reward *R* leads to more cooperation (Gill & Rosokha, 2020), and a higher temptation *T* leads to less cooperation. Dal Bó and Fréchette (2019) noted that when the reward is low, players are more likely to open with a defection on the first round. This indicates that when the gains from bilateral cooperation are not very high, players are more suspicious. They start by defecting even though they may still try to establish cooperation afterwards. In addition to these expected effects of payoffs, we note that there are many factors that can influence cooperation in games played with human subjects that are difficult to account for with standard evolutionary models. We already mentioned the effect of negative payoffs, as individuals are notoriously loss averse (Raub & Snijders, 1997). But also the value of the payoffs, which can be manipulated through the conversion rate into real money payments, can have an effect.

The last component that has a significant impact on behavior is the addition of noise through errors. Experimentally, errors can be implemented by having a choice be executed as its opposite. While any participant can find out that their own actions have been misimplemented, they have no way of finding out whether the co-player's choice was intentional. Instead, participants are only informed about the general rate with which errors occur. The presence of such noise changes how people play and which strategies they use. Already with a small level of noise, cooperation decreases (Aoyagi et al., 2003; Fudenberg et al., 2012; Li et al., 2022). This might be because strategies in treatments with noise tend to look further backward in time. That is, players condition not only on the previous round but also on older rounds. For example, players in Fudenberg et al. (2012) stated that they tended to give their co-player a benefit of the doubt. They would attribute the first defection to an error, and they would only start defecting themselves after the co-player defected multiple times. The authors describe this “leniency” and “forgiveness” as key components of strategies in noisy treatments.

4 | EVOLVING STRATEGIES

After having looked at the emerging cooperation rates, in the next step we wish to describe which strategies participants use. This endeavor, however, is non-trivial. After all, strategies are contingent plans—they tell a player what to do after any possible history of previous play. In contrast, in experiments participants often make decisions for one particular history, which makes it difficult to estimate how they would react to alternative scenarios.

In the literature, there have been different ways to deal with this problem. First, instead of asking participants to choose an action each round, we can ask them to choose their repeated-game strategies. Participants are informed that these elicited strategies are then used to determine how they act in the subsequent experiment. This is the so-called strategy method (Selten, 1967). Here, participants either chose from a menu of predefined strategies, or they define their memory-1 conditional strategies (i.e., for any outcome of the previous round, participants define with which probability they wish to cooperate in the next round, see Dal Bó & Fréchette, 2019; Gill & Rosokha, 2020). This method has the advantage that the results are clear; we can see which existing strategies are preferred. However, we lose a lot of nuance as individuals are usually more messy and hardly stick to one such strategy completely. By letting subjects choose from a finite strategy set, we also risk missing a strategy that would be popular had it existed in the menu of possible strategies. The other, more common approach is to infer strategy from actual choices (Dal Bó & Fréchette, 2011, 2018; Embrey et al., 2018; Friedman & Oprea, 2012; Fudenberg et al., 2012; Li et al., 2022; Mao et al., 2017; Milinski & Wedekind, 1998; Montero-Porras et al., 2022; Palfrey & Rosenthal, 1994; Wedekind & Milinski, 1996). This method is more cumbersome and different techniques exist. One is to use Bayesian inference to ask which strategy (out of a given set) is most likely to reproduce a participant's observed behavior. Because the set of possible strategies is determined by the researcher, this approach is subject to similar criticisms as the strategy method. The other option is to estimate conditional responses based on previous play (most often assuming that individuals react to the last round only). This approach, however, requires that participants in fact experience all possible game outcomes for which a response is to be estimated.

4.1 | Strategies in the finitely repeated prisoner's dilemma

After highlighting the difficulties that arise when estimating the participants' strategies, in the following we discuss which conclusions have been drawn with the above methods. In the finitely repeated prisoner's dilemma, conclusions are surprisingly clear. Here, the data suggests that a large fraction of participants can be accurately described by a particular class of conditional strategies. For a game of length n, these strategies define a threshold of rounds $k \leq n$ up to which they cooperate—unless the co-player defects before, in which case they defect for the remainder of the game (Embrey et al., 2018; Mao et al., 2017). These estimated strategies are in good agreement with previous models of cooperation in finitely repeated games (Kreps et al., 1982; McNamara et al., 2004).

4.2 | Indefinitely repeated prisoner's dilemma

In games in which there is always a probability of another encounter, results are more mixed. It seems that the dominating strategies are TFT (including some variants thereof), *ALLD* and *GRIM* (Breitmoser, 2015; Dal Bó & Fréchette, 2011, 2019, 2018; Fudenberg et al., 2012; Montero-Porras et al., 2022; Wedekind & Milinski, 1996). Papers that allow for longer memory either find that it is not necessary (Dal Bó & Fréchette, 2019), or that players simply prefer a more lenient version of *TFT*. This can be modulated by the payoffs chosen as demonstrated by Gill and Rosokha (2020): the higher the reward, the more lenient the strategies. One drawback of these results, especially when strategies are estimated from behavior, is that when players cooperate from beginning to the end of an interaction, there is no way to distinguish among several possible strategies. Adding noise can force a more diverse history of play, which makes it easier to tell strategies apart. When that is the case, longer-memory strategies and more lenient strategies become more popular (Fudenberg et al., 2012; Li et al., 2022). Surprisingly, however, these experiments give little support to *WSLS*, which usually emerges in evolutionary simulations (Hauert & Schuster, 1997; Nowak & Sigmund, 1993). These observations suggest that when evolutionary simulations predict cooperation to evolve based on *WSLS*, they might not reflect the true dynamic that underlies human cooperation. One aspect of *WSLS* that might be particularly counter-intuitive to human subjects is how it continues after a deviation from mutual cooperation (possibly because of an error). In that case, the strategy comes back to cooperation only after the interaction fell into mutual defection, not before. As such it is a little bit more forward looking than *GRIM* or *TFT*, which either never return to cooperation (*GRIM*), or only after the other player cooperated (*TFT*). This subtlety of *WSLS* might make it hard for humans to understand the true intentions of someone using this rule of behavior, even though it makes *WSLS* more robust to errors.

5 | MEMORY CONSTRAINTS

As discussed above, memory plays a key role in the strategies people play in a prisoner's dilemma. When strategies are constructed, we have a choice over how much memory we allow. Real humans are not so straightforward and simple. Already when estimating strategies, some papers limit memory by only considering memory-1 strategies. To some extent, there is a good argument to be made to limit memory, as many people would fail to remember exactly what happened in all previous rounds as length increases. In addition, in a real-life setting, interactions can span weeks, months, years, and people interact with many other interaction partners during that time-frame. All of this places some constraints on what can be realistically remembered of the details of the interaction.

A few papers have tested memory for cooperative actions explicitly. Stevens et al. (2011) and Winke and Stevens (2017) have the participants take part in a memory task where pictures of hypothetical partners as well as their action in a hypothetical game is displayed on screen. Treatments vary the number of total partners in a memory set or the number of “in-between” partners between two viewings of the same partner (Stevens et al., 2011). These studies suggest that overall, memory is extremely poor. Moreover, error rates further increase drastically as the number of "in-between" partners increases. As a consequence, when the researchers perform evolutionary simulations based on these error rates, the dominating strategies tend to be *ALLD* and *GRIM*. Interestingly, the total number of different partners does not influence memory. The authors conclude that traditional conditional strategies such as TFT are not realistic because in a setting with multiple partners, memory is not sufficiently accurate.

However, these studies look at memory without having subjects actually interact in a game. The results are quite different when participants must recognize and type hypothetical partners that they actually played a prisoner's dilemma with (computer partners with photographs). Volstorf et al. (2011) find that memory is highly accurate for recognition and categorization as cooperator or defector both immediately and 1 week after playing. The amount of cooperation with each type matches participants' memory performance: subjects cooperate less with partners that defected before. They also highlight that memory is best for rare types in the population, rather than best for defectors as previous literature has suggested. Similarly, Bell et al. (2017) find that memory for both defectors and cooperators is accurate when playing a repeated prisoner's dilemma against 16 different computer partners they encountered six times. These studies show that even if exact actions may not be remembered perfectly, human subjects have an accurate feeling of the kind of partner they are facing.

These papers tested memory for partners and their actions. Another aspect of memory is simply how a high load might affect cooperative behavior and strategies. Milinski and Wedekind (1998) investigated if strategy complexity is affected by memory load, which they find to be the case. When subjects play a memory game in parallel to the repeated prisoner's dilemma, they move from playing *WSLS* to *TFT*. However, the methods of this paper might not pass the test of time. Interactions were not anonymous and the entire lab saw the decisions of the participants. A more recent experiment by Duffy and Smith (2014) using the same distracting memory task finds that low load subjects are better able to condition their strategy on previous outcomes. Players in both the low and high load condition conditioned their strategies on previous actions, but only low load players seem to consider older actions.

These empirical results demonstrate that individuals remember the information they need in order to reciprocate cooperation. When the interaction is real, they are attentive to player types even when encountering dozens of multiples partners in one session and can remember these players accurately for days. At the same time, when the demands on memory are high, player tend to use simpler strategies but still maintain a similar level of cooperation. Sophisticated strategies that require long memory do not seem crucial to the emergence of reciprocity. Instead, simple rules of behavior relying on remembering types of players is often sufficient.

6 | BEYOND THE STANDARD PRISONER'S DILEMMA

In the previous sections, we restricted our attention to a particular class of experiments on reciprocity. In all cases, participants interacted with a fixed co-player in a prisoner's dilemma over a series of multiple (discrete) rounds. In the following, we briefly mention two natural extensions that highlight the particular flexibility that researchers have when conducting experiments with humans. One extension deals with cooperation in networked populations; the other extension explores how people cooperate when they make decisions in real time.

Most human interactions happen within a social network where individuals have relationships with many others. Abundant theoretical work suggests that such non-trivial interaction structures can have an impact on cooperation through the mechanism of network reciprocity (Allen et al., 2017; Nowak & May, 1992). This form of reciprocity argues that different network shapes and connectivity patterns allow players to cluster into cooperative groups. This natural occurring assortment makes cooperators less susceptible to exploitation. Several papers have tested this theory with human players in large to very large networks (Gracia-Lázaro et al., 2012; Grujic et al., 2010; Traulsen et al., 2010). Assuming that individuals have to choose the same action (cooperate or defect) against all their neighbors (as in the models), these studies find little evidence of clustering. Moreover, they find a similar decay in cooperation independent of the size and exact shape of the network, unless the benefit of cooperation is sufficiently large (Rand et al., 2014). As for strategies, a re-analysis of the three main papers found that players seem to ignore the payoffs of their neighbors when making decisions. Instead they simply chose their action based on how many cooperators are among their neighbors, as well as what they themselves did in the previous round (Grujic et al., 2014). These results highlight the importance of direct reciprocity, even when interacting with several connected players. However, when the number of interaction partners is more than just one, cooperation systematically decays, which is a common theme in multiplayer social dilemmas (Fischbacher and Gächter, 2010; Martinez-Martinez & Normann, 2022). Even when players use conditional cooperation and attempt to reciprocate, when the number of partners is too large, they struggle (Grujic et al., 2012).

However, most social networks are not static. Humans are usually able to end relationships with defectors and instead initialize interactions with other cooperators. Direct reciprocity in large networks of connected individuals can lead to cooperation if players can adjust their ties. To address this, Fehl et al. (2011) and Rand et al. (2011) investigate cooperation in dynamical networks. Here, players can cut their link to their neighbors when they are not satisfied with the relationship. Under this setup, cooperation is greatly enhanced as players learn to break ties with defectors. Moreover, this positive effect persists even if participants need to pay a substantial cost to cut ties (Bednarik et al., 2014). These results suggest that the mere possibility to quit an interaction is effective in promoting cooperation.

Dynamic networks can be realized in many different ways. Wang et al. (2012) allowed players to choose their new partner. Any new link had to be accepted by both parties and there was no upper limit on the number of connections of a player. The authors find that if the benefit of cooperative relationships is large enough compared to the cost of cooperating with a defector, players make the rational decision to create new cooperative ties rather than sever defective ones. Interestingly, this leads to a proliferation of defectors and lowers overall cooperation in the network. Antonioni et al. (2015) look at players' movements in a grid where they can choose their location relative to their neigbors. They find that cooperators do indeed cluster together. However, those cooperators at the boundaries get tired of being exploited by their defecting neigbors and start defecting, too. This leads cooperation to unravel. Nevertheless, in the vast majority of network experiments, subjects have to choose one action for all neighbors (see Fehl et al., 2011, for an exception). This design choice does not allow for proper one-to-one direct reciprocal relationships and does not treat the interactions as independent. The results of those studies could be very different if subjects were allowed to give targeted responses to each neighbor.

Another interesting variation on the classical prisoner's dilemma arises when people can make their decisions in real time (Friedman & Oprea, 2012; Martinez-Martinez & Normann, 2022). In corresponding experiments, players no longer make decisions in well-defined rounds. Rather they can choose with which action to start (cooperation or defection). After that, the game unfolds in continuous time, and people can revise their chosen action at any given point. Compared to the classical setup, this experimental design has several features that make it particularly attractive. For one, games tend to last shorter; players no longer need to make a sequence of decisions after which they need to be informed of the co-player's last decision. Rather decisions are made and information is provided in real time, such that supergames are typically finished in 1 or 2 min. At the same time, results from continuous-time experiments seem to be comparable to the classical setup. For example, for the finitely repeated prisoner's dilemma, Friedman and Oprea (2012) find that individual behavior is consistent with a conditional cutoff strategy. Participants cooperate until almost the end of the game, unless their opponent defected first, recovering similar results in the conventional repeated prisoner's dilemma (Embrey et al., 2018; Mao et al., 2017).

7 | DISCUSSION

Over the last decades, the repeated prisoner's dilemma has become the standard model for the evolution of direct reciprocity. It encapsulates the idea that individuals can maintain cooperation when they repeatedly interact in stable pairs, or small groups. By now, there is a rich theoretical literature that describes in which environments cooperation is to evolve, and which strategies are most effective in sustaining cooperation. In this article we compare these theoretical results to the empirical literature on human cooperation. Because the empirical literature on the prisoner's dilemma is vast, here we only present a selection of works. For a more comprehensive overview on the empirical literature, we recommend the invaluable resource of the cooperation databank (Spadaro et al., 2022), as well as other review articles (Dal-Bó & Fréchette, 2018; Romano et al., 2022). Perhaps somewhat surprisingly, our comparison shows that the predictive value of theoretical models is somewhat ambivalent. On the one hand, models seem to describe reasonably well for which parameters cooperation is most likely to evolve. In particular, the effect of parameter changes is often accurately predicted by these models: for example, increasing the expected length of the game does tend to increase cooperation; and similarly, increasing the payoffs for mutual cooperation makes people on average more cooperative.

On the other hand, models seem to be far less successful when it comes to predict the particular strategies that humans would use. For example, in indefinitely repeated games (those without a known end), evolutionary models often stress the success of strategies like *WSLS* (Hauert & Schuster, 1997; Kraines & Kraines, 1993; Martinez-Vaquero et al., 2012; Nowak & Sigmund, 1993; Stewart & Plotkin, 2014). In fact, this strategy has a number of appealing theoretical properties. It can resist invasion by unconditional cooperators, it is robust with respect to occasional errors and mistakes, and it is evolutionary stable when cooperation is sufficiently valuable (Hilbe et al., 2017). However, most empirical studies find little evidence for behavior consistent with *WSLS*, even in parameter regions in which this strategy is supposed to be strongly favored (e.g., Fudenberg et al., 2012; Li et al., 2022).

There is a number of reasons that might account for this mismatch. For example, evolutionary simulations are often run under rather restrictive parameters assumptions. Most importantly, many studies assume that mutations are rare, which allows researchers to simulate evolutionary processes more efficiently (Wu et al., 2012). When mutations are assumed to be rare, most populations tend to be monomorphic, which favors the evolution of equilibrium strategies like *WSLS*. On the other hand, data from experiments suggests that there is quite some variation in human behavior (Fudenberg et al., 2012; Li et al., 2022). In populations with many different strategies, more reciprocal strategies like *TFT* may have an advantage, because they are less prone to be exploited by any given opponent.

Another limitation of most evolutionary models is that people are often assumed to play each repeated game in isolation. In contrast, most human interactions do not happen in a strict uninterrupted sequence. Rather we engage in games with one individual at one time, only to interact with another group member a few minutes later. To date, there is little theoretical work that can describe how individuals keep an optimal record of their social interactions, and how they should react based on their record. While our discussion of memory constraints suggest that humans tend to remember the general nature of their co-player, there might be interesting interactions between the exact way how people memorize past interactions, and which strategies they use in response.

More generally, much of the previous research, both theoretically and experimentally, is restricted to constrained strategy sets. In particular, researchers often focus on memory-1 strategies, or on some of the simple strategies taken from the classic set described in Table 1. Even when more complex strategies are considered, they are typically longer memory extensions of essentially the same rules (for example, Fudenberg et al., 2012, considers nine variants of *TFT* out of a total set of 20 different strategies used in their analysis). Research might benefit from testing a more heterogenous set of strategies when investigating human behavior in the repeated prisoner's dilemma. These strategies should be explored in different environments, with different error rates and game lengths. Romero and Rosokha (2018) and Montero-Porras et al. (2022) specifically look at very long games and find that the way players punish, exploit, or forgive can be predicted by how long the interaction has the potential to last. Longer games and the presence of errors allow for richer behaviors and strategies, and could make for interesting future research into the dynamics of reciprocity.

AUTHOR CONTRIBUTIONS

Charlotte Rossetti: Conceptualization; writing—original draft; writing—review and editing. Christian Hilbe: Conceptualization; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

This work was supported by the European Research Council Starting Grant 850529: E-DIRECT, and by the Max Planck Society. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

This article does not contain any original data.

ORCID

Christian Hilbe https://orcid.org/0000-0001-5116-955X

REFERENCES

Akin, E. (2016). The iterated prisoner's dilemma: Good strategies and their dynamics. In I. Assani (Ed.), *Ergodic theory, advances in dynamics* (pp. 77–107). de Gruyter.

Allen, B., Lippner, G., Chen, Y.-T., Fotouhi, B., Momeni, N., Yau, S.-T., & Nowak, M. A. (2017). Evolutionary dynamics on any population structure. *Nature*, *544*, 227–230.

Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal*, *103*(418), 570–585.

Antonioni, A., Tomassini, M., & Sánchez, A. (2015). Short-range mobility and the evolution of cooperation: An experimental study. *Scientific Reports*, *5*(1), 10282.

Aoyagi, M., & Fréchette, G. (2009). Guillaume Fréchette. Collusion as public monitoring becomes noisy: Experimental evidence. *Journal of Economic Theory*, *144*(3), 1135–1165.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Bednarik, P., Fehl, K., & Semmann, D. (2014). Costs for switching partners reduce network dynamics but not cooperative behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20141661.

Bell, R., Koranyi, N., Buchner, A., & Rothermund, K. (2017). The implicit cognition of reciprocal exchange: Automatic retrieval of positive and negative experiences with partners in a prisoner's dilemma game. *Cognition and Emotion*, *31*(4), 657–670.

Blonski, M., Ockenfels, P., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, *3*, 164–192.

Boyd, R., & Lorberbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature*, *327*(6117), 58–59.

Breitmoser, Y. (2015). Cooperation, but no reciprocity: Individual strategies in the repeated prisoner's dilemma. *American Economic Review*, *105*(9), 2882–2910.

Colman, A. M. (2006). The puzzle of cooperation. *Nature*, *440*, 744–745.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*(2), 187–218.

Crandall, J. W., Oudah, M., Tennom, F., Ishowo-Oloko, S., Abdallah, J.-F., Bonnefon, M., Cebrian, A., Shariff, A., Goodrich, M. A., & Rahwan, I. (2018). Cooperating with machines. *Nature Communications*, *9*, 233.

Dal Bó, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review*, *95*, 1594–1604.

Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, *101*, 412–429.

Dal Bó, P., & Fréchette, G. R. (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*, *109*(11), 3929–3952.

Dal-Bó, P., & Fréchette, G. R. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, *56*, 60–114.

Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, *51*, 47–56.

Embrey, M., Fréchette, G. R., & Yuksel, S. (2018). Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*, *133*(1), 509–551.

Fehl, K., van der Post, D. J., & Semmann, D. (2011). Co-evolution of behaviour and social network structure promotes human cooperation. *Ecology Letters*, *14*(6), 546–551.

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, *100*(1), 541–556.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters*, *71*, 397–404.

Frank, M. R., Obradovich, N., Sun, L., Woon, W. L., LeVeck, B. L., & Rahwan, I. (2018). Detecting reciprocity at a global scale. *Science Advances*, *4*(1), eaao5348.

Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review*, *102*(1), 337–363.

Fudenberg, D., Dreber, A., & Rand, D. G. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, *102*, 720–749.

Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, *54*(3), 533–554.

Gill, D., & Rosokha, Y. (2020). Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma. SSRN 3652318.

Glynatsi, N., & Knight, V. (2020). A meta analysis of tournaments and an evaluation of performance in the iterated prisoner's dilemma. *arXiv* 2001.05911.

Glynatsi, N., & Knight, V. (2021). 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.

Gracia-Lázaro, C., Ferrer, A., Ruiz, G., Tarancón, A., Cuesta, J. A., Sánchez, A., & Moreno, Y. (2012). Heterogeneous networks do not promote cooperation when humans play a prisoner's dilemma. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(32), 12922–12926.

Grujic, J., Eke, B., Cabrales, A., Cuesta, J. A., & Sánchez, A. (2012). Three is a crowd in iterated prisoner's dilemmas: Experimental evidence on reciprocal behavior. *Scientific Reports*, *2*(1), 1–7.

Grujic, J., Fosco, C., Araujo, L., Cuesta, J., & Sanchez, A. (2010). Social experiments in the mesoscale: Humans playing a spatial prisoner's dilemma. *PLoS One*, *5*, e13749.

Grujic, J., Gracia-Lázaro, C., Milinski, M., Semmann, D., Traulsen, A., Cuesta, J. A., Moreno, Y., & Sánchez, A. (2014). A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Scientific Reports*, *4*, 4615.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Hauert, C., & Schuster, H. G. (1997). Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: A numerical approach. *Proceedings of the Royal Society B*, *264*, 513–519.

Hilbe, C., Chatterjee, K., & Nowak, M. A. (2018). Partners and rivals in direct reciprocity. *Nature Human Behaviour*, *2*(7), 469–477.

Hilbe, C., Martinez-Vaquero, L. A., Chatterjee, K., & Nowak, M. A. (2017). Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences United States of America*, *114*, 4715–4720.

Hilbe, C., Röhl, T., & Milinski, M. (2014). Extortion subdues human players but is finally punished in the prisoner's dilemma. *Nature Communications*, *5*, 3976.

Hilbe, C., Traulsen, A., & Sigmund, K. (2015). Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games and Economic Behavior*, *92*, 41–52.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *4*, 399–425.

Knight, V., Campbell, O., Harper, M., Langner, K. M., Campbell, J., Campbell, T., Carney, A., Chorley, M., Davidson-Pilon, C., Glass, K., Glynatsi, N., Ehrlich, T., Jones, M., Koutsovoulos, G., Tibble, H., Muller, J., Palmer, G., Petnunov, P., Slavin, P., … Molden, K. (2016). An open framework for the reproducible study of the iterated prisoner's dilemma. *Journal of Open Research Software*, *4*, e35.

Kraines, D. P., & Kraines, V. Y. (1993). Learning to cooperate with pavlov an adaptive strategy for the iterated prisoner's dilemma with noise. *Theory and Decision*, *35*, 107–150.

Kreps, D., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory*, *27*, 245–252.

Li, J., Zhao, X., Li, B., Rossetti, C., Hilbe, C., & Xia, H. (2022). Evolution of cooperation through cumulative reciprocity. *Nature Computational Science*, *2*, 677–686.

Mailath, G. J., & Samuelson, L. (2006). *Repeated games and reputations*. Oxford University Press.

Mao, A., Dworkin, L., Suri, S., & Watts, D. J. (2017). Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoner's dilemma. *Nature Communications*, *8*, 13800.

Martinez-Martinez, I., & Normann, H. (2022). Cooperation in multiplayer dilemmas. *SSRN*. https://doi.org/10.2139/ssrn.4294660

Martinez-Vaquero, L. A., Cuesta, J. A., & Sanchez, A. (2012). Generosity pays in the presence of direct reciprocity: A comprehensive study of 2×2 repeated games. *PLoS One*, *7*(4), E35135.

McNamara, J. M., Barta, Z., & Houston, A. I. (2004). Variation in behaviour promotes cooperation in the prisoner's dilemma game. *Nature*, *428*, 745–748.

Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences United Staes of America*, *95*, 13755–13758.

Montero-Porras, E., Grujic, J., Fernández Domingos, E., & Lenaerts, T. (2022). Inferring strategies from observations in long iterated prisoner's dilemma experiments. *Scientific Reports*, *12*, 7589.

Murase, Y., & Baek, S. K. (2020). Five rules for friendly rivalry in direct reciprocity. *Scientific Reports*, *10*, 16904.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*, 1560–1563.

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, *359*, 826–829.

Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, *364*, 56–58.

Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups* (revised ed.). Harvard University Press.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Palfrey, T. R., & Rosenthal, H. (1994). Repeated play, cooperation and coordination: An experimental study. *The Review of Economic Studies*, *61*(3), 545–565.

Press, W. H., & Dyson, F. D. (2012). Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 10409–10413.

Raihani, N. J. (2021). *The social instinct—How cooperation shaped the world*. Random House.

Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences United States of America*, *108*, 19193–19198.

Rand, D. G., Nowak, M. A., Fowler, J. H., & Christakis, N. A. (2014). Static network structure can stabilize human cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 17093–17098.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma*. University of Michigan Press.

Rapoport, A., Seale, D. A., & Colman, A. M. (2015). Is tit-for-tat the answer? On the conclusions drawn from axelrod's tournaments. *PLoS One*, *10*(7), e0134128.

Raub, W., & Snijders, C. (1997). Gains, losses, and cooperation in social dilemmas and collective action: The effects of risk preferences. *Journal of Mathematical Sociology*, *22*(3), 263–302.

Reiter, J. G., Hilbe, C., Rand, D. G., Chatterjee, K., & Nowak, M. A. (2018). Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nature Communications*, *9*, 555.

Romano, A., Saral, A. S., & Wu, J. (2022). Direct and indirect reciprocity among individuals and groups. *Current Opinion in Psychology*, *43*, 254–259.

Romero, J., & Rosokha, Y. (2018). Constructing strategies in the indefinitely repeated prisoner?s dilemma game. *European Economic Review*, *104*, 185–219.

Rossetti, C., Hilbe, C., & Hauser, O. P. (2022). (Mis)perceiving cooperativeness. *Current Opinion in Psychology*, *43*, 151–155.

Roth, A. E., & Murnighan, J. K. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology*, *17*(2), 189–198.

Schmid, L., Chatterjee, K., Hilbe, C., & Nowak, M. (2021). A unified framework of direct and indirect reciprocity. *Nature Human Behaviour*, *5*, 1292–1302.

Schweinfurth, M. K., & Taborsky, M. (2018). Reciprocal trading of different commodities in Norway rats. *Current Biology*, *28*, 594–599.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In H. Sauermann (Ed.), *Beitra¨ge zur experimentellen Wirtschaftsforschung* (pp. 136–168). J.C.B Mohr (Siebeck).

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, *4*, 25–55.

Sigmund, K. (2010). *The calculus of selfishness*. Princeton University Press.

Spadaro, G., Tiddi, I., Columbus, S., Jin, S., Ten Teije, A., CoDa Team, & Baillet, D. (2022). The cooperation databank: Machine-readable science accelerates research synthesis. *Perspectives on Psychological Science*, *17*, 1472–1489.

Stevens, J. R., Volstorf, J., Schooler, L. J., & Rieskamp, J. (2011). Forgetting constrains the emergence of cooperative decision strategies. *Frontiers in Psychology*, *1*, 235.

Stewart, A. J., & Plotkin, J. B. (2014). Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences United States of America*, *111*(49), 17558–17563.

Stewart, A. J., & Plotkin, J. B. (2015). The evolvability of cooperation under local and non-local mutations. *Games*, *6*(3), 231–250.

Szolnoki, A., & Perc, M. (2014). Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Scientific Reports*, *4*, 5496.

Tazelaar, M. J. A., Van Lange, P. A. M., & Ouwerkerk, J. W. (2004). How to cope with "noise" in social dilemmas: The benefits of communication. *Journal of Personality and Social Psychology*, *87*(6), 845–859.

Tkadlec, J., Hilbe, C., & Nowak, M. A. (2023). Mutation enhances cooperation in direct reciprocity. *Proceedings of the National Academy of Sciences United States of America*, *120*(20), e2221080120.

Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H.-J., & Milinski, M. (2010). Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences United States of America*, *107*, 2962–2966.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*, 35–57.

Van Lange, P. A. M., Ouwerkerk, J. W., & Tazelaar, M. J. A. (2002). How to overcome the detrimental effects of noise in social interaction: The benefits of generosity. *Journal of Personality and Social Psychology*, *82*(768–780), 768–780.

van Veelen, M., García, J., Rand, D. G., & Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences United States of America*, *109*, 9929–9934.

Volstorf, J., Rieskamp, J., & Stevens, J. R. (2011). The good, the bad, and the rare: Memory for partners in social interactions. *PLoS One*, *6*(4), e18945.

Wang, J., Suri, S., & Watts, D. J. (2012). Cooperation and assortativity with dynamic partner updating. *Proceedings of theNational Academy of Sciences of the United States of America*, *109*(36), 14363–14368.

Wedekind, C., & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(7), 2686–2689.

Winke, T., & Stevens, J. R. (2017). Is cooperative memory special? The role of costly errors, context, and social network size when remembering cooperative actions. *Frontiers in Robotics and AI*, *4*, 52.

Wörle, M., & Paulus, M. (2019). Normative foundations of reciprocity in preschoolers. *Journal of Experimental Child Psychology*, *188*, 104693.

Wu, B., Gokhale, C. S., Wang, L., & Traulsen, A. (2012). How small are small mutation rates? *Journal of Mathematical Biology*, *64*, 803–827.

Xu, B., Zhou, Y., Lien, J. W., Zheng, J., & Wang, Z. (2016). Extortion can outperform generosity in iterated prisoner's dilemma. *Nature Communications*, *7*, 11125.

Zhang, H. (2018). Errors can increase cooperation in finite populations. *Games and Economic Behavior*, *107*, 203–219.

How to cite this article: Rossetti, C. S. L., & Hilbe, C. (2023). Direct reciprocity among humans. *Ethology*, *00*, 1–13. https://doi.org/10.1111/eth.13407