

A unified framework of direct and indirect reciprocity

Laura Schmid^{1,*}, Krishnendu Chatterjee¹, Christian Hilbe^{2,†}, Martin A. Nowak^{3,†}

¹IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

²Max Planck Research Group Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, 24306 Ploen, Germany

³ Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA

† These authors contributed equally

*E-mail: laura.schmid@ist.ac.at

Direct and indirect reciprocity are key mechanisms for evolution of cooperation. Direct reciprocity means individuals use their own experience to decide whether to cooperate with another person. Indirect reciprocity means they also consider the experiences of others. Although the two mechanisms are intertwined, they are typically studied in isolation. Here, we introduce a mathematical framework that allows us to explore both kinds of reciprocity simultaneously. We show that the well-known strategy ‘Generous Tit-for-Tat’ of direct reciprocity has a natural analogue in indirect reciprocity, which we call ‘Generous Scoring’. With an equilibrium analysis, we characterize under which conditions either of the two strategies can maintain cooperation. With simulations, we additionally explore which kind of reciprocity evolves when members of a population engage in social learning to adapt to their environment. We find that indirect reciprocity evolves if any two individuals meet only occasionally, when information about others is reliable, and when strategy mutations are rare. Our results draw unexpected connections between direct and indirect reciprocity, while highlighting important differences regarding their evolvability.

Reciprocity is a principle that guides many aspects of our social life¹⁻⁴. Whenever people repay a favor, write a positive evaluation of an online seller, or build up trust over multiple interactions, they engage in reciprocal behavior. Previous work distinguishes two kinds of reciprocity. Direct reciprocity⁵⁻¹⁹ means that my behavior towards you depends on what you have done to me. Indirect reciprocity²⁰⁻²⁴ means that my behavior towards you also depends on what you have done to others. Direct reciprocity requires that the same individuals interact repeatedly, which enables them to respond to their interaction partner in future transactions (**Fig. 1a**). Indirect reciprocity does not require individuals to have a joint history of previous interactions, nor does it require them to ever meet again. It is solely based on the premise that by helping someone, you can increase your public standing. This reputational gain is valuable in future interactions with others (**Fig. 1b**). Experiments suggest that human behavior is shaped by both direct^{25,26} and indirect reciprocity²⁷⁻²⁹.

While direct and indirect reciprocity are related, the respective models are strikingly different. Studies of direct reciprocity⁵⁻¹⁹ report substantial cooperation rates even if subjects only remember a minimum of information. Successful strategies like Tit-for-Tat⁵ (*TFT*) and Generous Tit-for-Tat⁶ (*GTFT*) only keep track of the very last interaction. In contrast, studies of indirect reciprocity stress that cooperation can only be maintained when strategies are sufficiently complex²¹⁻²³. To describe how complex strategies need to be, this literature distinguishes different classes of strategies. The most elementary class are the first-order strategies, where a player's reputation only depends on her previous actions. A well-known example is image scoring²⁰. Here, reputations are represented by an integer score. A player's score increases when she cooperates, and it drops when she defects. Individuals only cooperate with those who have a sufficiently high score. Classic image scoring, however, is unstable²¹. After all, individuals have no incentive to retaliate against defectors, because this would impede their own score. This instability suggests considering second-order strategies. Here, reputations do not only depend on what an individual did, but also to whom. For example, when an individual defects against a co-player with a bad reputation, this defection may be considered as justified. The hierarchy of strategies can be further extended to third order. Here, players additionally take into account the focal individual's reputation.

In a landmark study, Ohtsuki & Iwasa explored which strategies of up to third order sustain cooperation²². In their study, reputations are required to be binary (good or bad), strategies are deterministic (the same behaviour always yields the same reputation), and all information is public and mutually agreed upon. Within this setup, they show there are no stable first-order strategies that give rise to cooperation. However, there are two second-order strategies and six third-order strategies. These so-called leading eight strategies can sustain cooperation because they allow for more sensible judgments than image scoring. At the same time, they require more information than most well-known strategies of direct reciprocity.

The two kinds of reciprocity also differ in how susceptible they are to misunderstandings and other types of errors. Whereas *GTFT* and similar strategies of direct reciprocity are robust with respect to noise^{4,30}, the leading eight strategies of indirect reciprocity are not^{31,32}.

Due to such differences, it has been difficult to analyse the two modes of reciprocity within a single theoretical framework. Previous work has taken two approaches. The first approach is to suggest particular

strategies that combine elements of direct and indirect reciprocity, and to analyse their stability^{33,34}. The second approach uses computer simulations to let different strategies compete^{35–37}. Two noteworthy studies of the latter kind are by Nakamaru & Kawata³⁶ and Seki & Nakamaru³⁷. They explore the evolution of reciprocity when players can fake their own reputation, or misrepresent the reputation of others. The two studies observe that when outside information becomes unreliable, players tend to ignore it. Computational studies, however, make it difficult to compare the different kinds of reciprocity directly. They often involve comparisons between strategies of different complexity. Moreover, the relative advantage of each type of reciprocity can only be inferred by comparing simulations for specific parameter choices.

Instead, here we propose a framework that can be analysed explicitly. For our study, we extend the theory of zero-determinant strategies from direct^{8–16} to indirect reciprocity. This approach allows us to draw analytic conclusions when comparing the two mechanisms.

Results

A unified framework of reciprocity.

To introduce a model that entails both kinds of reciprocity, we consider a population of n players. Players engage in the following sequence of interactions. In the beginning, two players are randomly drawn from the population to interact in one round of the prisoner’s dilemma. In this game, each player independently decides whether to cooperate (C) or defect (D). Cooperation means to pay a cost $c > 0$ to provide a benefit $b > c$ for the co-player. After each such interaction, with probability d again two players are randomly drawn from the population to engage in another round of the game. Otherwise, with probability $1 - d$, no further interaction occurs. Once there are no more interactions, we calculate the payoffs of each player by averaging over all pairwise games in which the respective player participated.

To make their decisions, players represent each co-player by a separate finite-state automaton. Each automaton has two possible states, labeled as Good (G) and Bad (B), see **Fig. 1c**. Players cooperate with those co-players they currently deem as good, and they defect against those they consider as bad. They update the current state of each co-player according to their strategy (y, p, q, λ) . The parameter y is the initial probability that a co-player is considered as good, in the absence of any information (**Fig. 1d**). The parameters p (and q), determine the probability to assign a good reputation to a co-player who has just cooperated (defected) in a direct interaction (**Fig. 1e,f**). The parameter λ is a player’s receptivity to indirect information. If a co-player interacts with a third party, then with probability λ the focal player updates that co-player’s state accordingly. In that case, again the co-player obtains a good reputation with probability p or q , depending on whether she cooperated or defected (**Fig. 1g, Extended Data Fig. 1**). For simplicity, we assume in the main text that all population members observe everybody else’s interactions. However, they may misinterpret the outcome of games between others with probability ε . When such an observation error occurs, a third-party observer mistakenly interprets a player’s cooperation as defection, and vice versa.

If $\lambda = 0$ for all individuals, players base their decisions entirely on their own experiences. In that case

our framework reduces to the standard model of direct reciprocity with reactive strategies⁶. On the other hand, if $\lambda = 1$ for all individuals, then players take all interactions of their opponents equally into account, no matter whether they are directly involved. In that case our framework yields a model of indirect reciprocity among players with stochastic first-order strategies³⁸. It is important to note that even players with $\lambda = 1$ do not ignore any directly obtained information they may have. For example, if the same two individuals are chosen to interact for two consecutive rounds, their second-round behavior will naturally depend on the outcome of the first round. However, in large populations in which such consecutive encounters are unlikely, the role of direct information on players with $\lambda = 1$ becomes negligible. In **SI Section 6.2**, we compare this baseline model with an alternative setup, where we consider a ‘purified’ version of indirect reciprocity. In that alternative setup, players can choose to ignore all direct experiences they have, such that they solely rely on third-party information. The results of that alternative model are similar to the results presented herein.

Our strategy space contains several well-known strategies of direct and indirect reciprocity. Examples include $TFT = (1, 1, 0, 0)$, $GTFT = (1, 1, q, 0)$, and an elementary image scoring rule³⁹ referred to as Simple Scoring⁴⁰, $SCO = (1, 1, 0, 1)$. However, our model is more general than these previous studies on either direct or indirect reciprocity in two ways. First, it allows for populations in which some players use direct reciprocity ($\lambda = 0$), whereas others use indirect reciprocity ($\lambda = 1$). Second, it allows players to combine the two modes of reciprocity, by choosing $0 < \lambda < 1$. In that case, players always take direct experiences into account, but they would occasionally also consider a co-player’s interactions with others.

Models of indirect reciprocity often assume ‘public information’^{22,23}. This does not only mean that all individuals learn all relevant information. Instead, respective models also assume that everyone agrees on each co-player’s reputation. Such an assumption can be problematic when individuals receive information from independent sources, or when information transmission is noisy³². Moreover, even if individuals agree on all past events, they may still disagree on which reputation a co-player should have if they apply different social norms. Such different assessments can easily arise, for example, when some players base their decisions on direct reciprocity, whereas others use indirect reciprocity. Because we are exactly interested in such scenarios, our model is necessarily one of ‘private information’, as in Nakamaru & Kawata³⁶ and Seki & Nakamaru³⁷. As a result, different players may hold different views on any given population member.

Throughout the main text, we will use the above baseline framework to explore the dynamics of direct and indirect reciprocity. However, in **SI Section 6**, we explore the effect of several model extensions. In particular, we discuss how our results change when we allow for alternative kinds of errors^{41–43} and for incomplete information⁴⁴. Moreover, we describe how our framework can be adapted to capture more complex strategies, including finite-state automata with more than two states^{18,45} or the leading eight²².

Equilibrium conditions for reciprocal cooperation.

Because the strategies of the baseline model only require first-order information, we can compute the players’ payoffs explicitly. The respective formula, derived in detail in **SI Section 3**, is valid for any population

size, arbitrary population compositions, and all parameter values. Based on this explicit representation of payoffs, we first characterised all Nash equilibria among the strategies (y, p, q, λ) . In a Nash equilibrium, no player can improve her payoff by unilaterally deviating (not even by using a more complex strategy that uses arbitrary amounts of past information). By extending the theory of zero-determinant strategies^{8–12}, we find that for every $\lambda \in [0, 1]$, there can be exactly one generic Nash equilibrium strategy (y, p, q, λ) that yields full cooperation. These strategies are explicitly derived in **SI Section 4**. In the following, we summarise the corresponding results.

For direct reciprocity ($\lambda = 0$) the unique strategy that yields stable cooperation is given by the classical *GTFT* strategy (**Fig. 2a**), with $y = p = 1$ and

$$q_0 = 1 - \frac{c}{\delta b}. \quad (1)$$

Here, δ is the probability that two interacting players interact again some time in the future. This pairwise continuation probability can be derived from the population-wide continuation probability d (**SI Section 4**). For indirect reciprocity ($\lambda = 1$), the Nash equilibrium has $y = p = 1$ and

$$q_1 = 1 - \frac{1 + (n-2)\delta}{1 + (n-2)(1-2\varepsilon)} \frac{c}{\delta b}. \quad (2)$$

In analogy to *GTFT*, we call this strategy Generous Scoring (*GSCO*, **Fig. 2b**). Both strategies have in common that they always assign a good reputation to cooperators, and that they occasionally assign a good reputation to defectors. However, they differ in which information they take into account when making these assessments. While *GTFT* only considers direct interactions, *GSCO* takes all interactions of a co-player into account.

The above descriptions of *GTFT* and *GSCO* only give rise to a sensible strategy if their q is non-negative. By requiring $q \geq 0$, equations (1) and (2) thus characterise when cooperation can be sustained at all. We find that the game's continuation probability needs to be sufficiently large, $\delta \geq \delta_\lambda$. The respective threshold values for direct ($\lambda = 0$) and indirect ($\lambda = 1$) reciprocity are

$$\delta_0 = \frac{c}{b} \quad \text{and} \quad \delta_1 = \frac{c}{b + (n-2)((1-2\varepsilon)b - c)}. \quad (3)$$

The threshold δ_0 for direct reciprocity is simply given by the cost-to-benefit ratio of cooperation³. The threshold δ_1 for indirect reciprocity can be greater or lower, depending on whether or not outside information is sufficiently reliable (i.e., depending on whether the probability ε of an observation error is greater or lower than $(1 - c/b)/2$). The two thresholds in (3) give rise to four possible cases (**Fig. 2c**). Either (i) cooperation is not feasible at all, (ii) it is only feasible through indirect reciprocity, (iii) it is only feasible through direct reciprocity, or (iv) it is feasible through both kinds of reciprocity. We have derived analogous thresholds for δ under the alternative assumption that both direct and indirect observations are subject to the same error rate (**SI Section 6.1**). In that case, the third region vanishes: if cooperation is feasible at all, it is always feasible through indirect reciprocity (**Fig. 2d**).

In addition to the extremal cases of direct reciprocity ($\lambda = 0$) and indirect reciprocity ($\lambda = 1$), we have also explored whether the equilibrium conditions for a cooperative equilibrium can be met more easily if players use intermediate values of λ . Interestingly, the answer is negative. Specifically, we prove that if there is a cooperative Nash equilibrium for some $0 < \lambda < 1$, then either *GTFT* or *GSCO* is already an equilibrium. From an equilibrium perspective, intermediate degrees of receptivity thus do not further extend the possibilities for cooperation. Moreover, in the limit of rare errors we find that the conditions (3) are strict even as we allow for arbitrarily complex strategies: if neither *GTFT* nor *GSCO* can sustain cooperation for the given parameters of the game, no other Nash equilibrium can.

Comparing the dynamics of direct and indirect reciprocity.

The previous equilibrium results highlight different strategies that can maintain cooperation if adopted by sufficiently many in the population. However, the above results do not imply that such strategies would automatically evolve. After all, $ALLD = (0, 0, 0, \lambda)$ is also an equilibrium for all parameter values (**SI Section 4**). In a next step, we have thus explored under which conditions cooperation can emerge when players engage in social learning.

To this end, we no longer assume that players use equilibrium strategies. Rather they may start out with some arbitrary strategy (y, p, q, λ) . Over time players adopt new strategies based on a pairwise comparison process^{46,47}. This process assumes that in each time step, one individual is randomly drawn from the population. This player then has the opportunity to revise her strategy. She can do so by either adopting a randomly chosen strategy (akin to a mutation in biological models), or by imitating the strategy of another group member (akin to selection). Imitation events are biased such that strategies with a high payoff have a better chance to be imitated (see **Methods**). This elementary strategy updating step is then iterated over many time periods. We use simulations to record which strategies the players adopt over time and how often they cooperate. To this end, we sometimes assume that mutations are rare. The limit of rare mutations is mathematically well understood^{48–52} and it has been prominently employed in previous studies of reciprocity^{53–59} and beyond^{60–64}. When mutations are rare, the population consists of at most two strategies, residents and mutants. The mutant strategy goes extinct or fixes before the next mutation arises. The assumption of rare mutations allows simulations to be run more efficiently. This in turn makes it easier to systematically explore the entire strategy space (see **SI Section 5** for details). We complement the respective results with simulations with frequent mutations.

We first explore the two limiting cases of reciprocity separately, by fixing either $\lambda = 0$ or $\lambda = 1$. We consider two different scenarios (**Fig. 3**). In the first scenario, individuals interact only for a few rounds. In the other scenario, we consider the limiting case that they interact for infinitely many rounds. This limit has been employed in many previous studies^{6–12} as it naturally reduces the dimension of the strategy space (see **Fig. 3a,b**). Similar results can be obtained if the number of rounds is large but finite (**Fig. 4**). In all scenarios we observe that for rare mutations, players either tend to adopt a strategy close to always defect, $ALLD = (0, 0, 0, \lambda)$, or a conditionally cooperative strategy $(1, 1, q, \lambda)$. As expected from our equilibrium analysis, indirect reciprocity is overall more favourable to cooperation when individuals interact only for

a few rounds. Interestingly, however, direct reciprocity is more effective in maintaining cooperation when many rounds are played, even in the absence of any observation errors.

To gain some analytical understanding for why direct reciprocity becomes superior, we consider an initial population that either employs *ALLD* or a conditionally cooperative strategy. For both resident strategies, we record how long it takes until a different strategy can invade (**Fig. 3c,d**) and which strategies are most likely to do so (**Fig. 3e,f**). When many games are being played, conditional cooperators have a similar invasion time for both direct and indirect reciprocity (**Fig. 3d**). However, *ALLD* can be invaded more easily when players use direct reciprocity. To explore this differential robustness of defectors, we analysed the competition between *ALLD* and a conditionally cooperative strategy $(1, 1, q, \lambda)$. When only these two strategies are present, the respective payoffs π_D and π_C can be calculated explicitly (see **SI Section 5**). In the limit of large populations and rare errors, the payoffs under direct reciprocity ($\lambda = 0$) become

$$\begin{aligned}\pi_C^0 &= (b-c) \cdot z - (1-\delta+\delta q)c \cdot (1-z) \\ \pi_D^0 &= (1-\delta+\delta q)b \cdot z.\end{aligned}\tag{4}$$

Here, z is the fraction of conditional cooperators in the population. Eq. (4) yields two insights. First, provided that $q < 1 - c/(\delta b)$, the dynamics is bistable. If cooperators are common ($z \approx 1$), they have the higher payoff. In contrast, when cooperators are rare ($z \approx 0$), defectors are favoured. Second, the payoff of the two strategies increases linearly in the fraction of cooperators. When we perform the same analysis for indirect reciprocity ($\lambda = 1$), we obtain

$$\begin{aligned}\pi_C^1 &= \frac{q + q(1-q)(1-z)}{1 - (1-q)z} \cdot z(b-c) - q(1-z)c \\ \pi_D^1 &= qb \cdot z.\end{aligned}\tag{5}$$

Again, for q sufficiently small these payoffs result in a bistable competition. However, while the defectors' payoffs continue to increase linearly in the fraction of cooperators, the cooperators' payoffs are now nonlinear (**Fig. 3g,h**).

This analysis highlights two crucial effects that distinguish indirect from direct reciprocity. On the one hand, indirect reciprocity leads to a faster spread of information throughout a population. As a consequence, indirect reciprocity is more effective in restricting the payoff of a defector (i.e., $\pi_D^1 < \pi_D^0$ for all $z > 0$). On the other hand, successful cooperation in indirect reciprocity is based on non-linear synergy effects. Cooperators only obtain high payoffs when they are sufficiently common. Which of the two effects is dominant depends on the population size, the error rate, and on how often players interact on average (**Extended Data Fig. 2**). Once players interact for many rounds, indirect reciprocity ceases to have any advantage (because $\pi_D^1 = \pi_D^0$ for $\delta \rightarrow 1$). In that case, defectors are always more readily invaded under direct reciprocity.

Due to their nonlinear returns, cooperative strategies of indirect reciprocity are most effective when they are common. This observation suggests that indirect reciprocity may be less likely to evolve when the evolutionary process itself prevents cooperative strategies to form a large majority. Such a case can occur,

for example, when mutations are abundant, such that many different strategies are routinely present in the population. To explore this issue in more detail, we have systematically varied the mutation rate of the evolutionary process (**Fig. 4**). Indeed, while mutations only have a minor effect on direct reciprocity, the impact on indirect reciprocity can be substantial as mutation rates become large. In particular, around $\mu n = 1$ (i.e., once there is more than one mutation per generation), cooperation rates decline quickly. Further simulations suggest that this downfall of cooperation is due to both a reduced stability of conditionally cooperative strategies and an enhanced stability of populations with a majority of defectors (**Fig. 4g-i**). These results highlight an important difference between direct and indirect reciprocity. While reactive strategies of direct reciprocity are largely robust to mutations, the corresponding strategies of indirect reciprocity are more sensitive. Strategies like Generous Scoring are most powerful in environments with little noise. To spread, they do not only need outside information to be faithful (small ε) but also the evolutionary process (small μ).

The co-evolution of direct and indirect reciprocity.

The above findings raise the question whether the players themselves are able to learn when to use indirect information. To explore this issue, we have first considered a simplified setup in which players can freely choose between all strategies (y, p, q, λ) where either $\lambda = 0$ and $\lambda = 1$. That is, players can choose whether they only take direct interactions into account, or whether they take all of a co-player's interactions equally into account. We study three different scenarios in the limit of rare mutations: one with noisy information and few interactions (**Fig. 5a**), one with reliable information and intermediately many interactions (**Fig. 5b**), and one with noisy information and many interactions (**Fig. 5c**). The results confirm our previous analytical findings. While defectors are predominant in the first scenario, individuals adopt conditionally cooperative strategies in the second and third scenario, showing a bias towards indirect and direct reciprocity, respectively. In a next step, we have systematically varied how often individuals interact with each other, and how noisy third-party information is (**Fig. 5d**). Again, we find that indirect reciprocity is most abundant when there are intermediately many rounds, such that cooperation cannot evolve through direct reciprocity alone.

We have repeated all simulations for an evolutionary process with more frequent mutations (**Fig. 5e-h**). While the qualitative results are similar, we recover our previous observation that larger mutation rates disfavour indirect reciprocity. Even in those parameter regions in which individuals learn to incorporate third-party information, evolving cooperation rates tend to be lower than in the scenario with rare mutations (**Fig. 5f**). The effect of other game parameters on the evolution of cooperation is discussed in **SI Section 5**, and visualised in **Extended Data Fig. 3** and **Extended Data Fig. 4**.

Finally, we have also explored which strategies evolve when players can adopt intermediate values of λ (**Extended Data Fig. 12**). To allow for a fair comparison between direct and indirect reciprocity, mutant strategies are drawn such that an average mutant would resort to their direct experience in approximately half of the cases (for details, see **SI Section 5.4**). Overall, we observe a similar trend as before: When information is noisy and there are very few rounds, individuals learn not to cooperate (**Extended Data**

Fig. 12e); when there is little noise and intermediately many interactions, individuals learn to cooperate predominantly based on indirect information (**Extended Data Fig. 12f**); and when there is an intermediate amount of noise and many interactions, individuals tend to cooperate based on direct information (**Extended Data Fig. 12g**).

Discussion

When deciding whether to cooperate, humans often resort to the co-player's reputation arising from third party interactions^{27,28}, sometimes even if the two players have a joint history of direct interactions²⁹. Most theoretical studies, however, do not investigate how subjects choose between these two sources of information. They either study direct reciprocity using repeated games⁵⁻¹⁸ or indirect reciprocity using donor-recipient games²⁰⁻²³. Here, we have proposed a general framework that unifies direct and indirect reciprocity.

To make such a comparison between different kinds of reciprocity most transparent, throughout the main text we have focussed on a comparably simple setup. For example, we have not modelled explicitly how information from third party interactions spreads throughout a population. We have assumed that individuals observe each others' interactions directly. Instead, one may equally assume that individuals use rumours and gossip to exchange information about their past experiences with other population members. Such communication can add another layer of complexity to the model because players may have an incentive to strategically misrepresent their reputation. For example, defectors are naturally incentivised to prevent others from faithfully learning about their past behaviours. As demonstrated in previous work^{36,37}, such miscommunication does not render cooperation impossible. However, it imposes additional bounds on when indirect reciprocity can evolve. While our model does not consider the effects of false gossip explicitly, we may capture some of its workings by assuming that observation errors may be biased. For example, acts of defection may be more likely to be misperceived than acts of cooperation. In **SI Section 6**, we show that all our analytical results naturally carry over to this more realistic setting.

Similarly, while we have explored the effect of observation errors in detail, we have neglected other types of errors. As an example, implementation errors have received considerable attention in the previous literature⁶⁵. Such errors lead players to mis-execute their intended actions. They may fail to cooperate although they planned to do so, perhaps because of a 'trembling hand'⁶⁶. The consequences of implementation errors can be rather different from observation errors, because only the former become publicly known. Nevertheless, we can show in **SI Section 6** that such implementation errors can be naturally included into our framework (see also **Extended Data Fig. 5**).

Finally, in the main text we have restricted ourselves to the simplest class of strategies, which only depend on a player's previous action. Within this class, we have identified a remarkable new strategy of indirect reciprocity. This strategy, called Generous Scoring (*GSCO*), is the analogue of Generous Tit-for-Tat (*GTFT*)⁶. It routinely cooperates with other cooperators, but it is also willing to forgive a defector occasionally. Unlike *GTFT*, however, *GSCO* does not require repeated interactions between two players; it can sustain cooperation even if individuals are likely to never meet again. When previous research on

indirect reciprocity identified stable cooperative strategies, the strategies are only shown to be stable within a given strategy class^{22,23}. This kind of analysis does not rule out that the respective equilibria turn out to be unstable once more complex mutant strategies are permitted. In contrast, Generous Scoring is a Nash equilibrium with respect to all possible mutant strategies, independent of whether mutants themselves use direct or indirect reciprocity, or how much information they are able to process.

This stability of Generous Scoring may be surprising. After all, first-order strategies have been suspected to be incapable of sustaining cooperation^{21–23}. For example, Image Scoring is unstable because players have no incentive to retaliate against defectors in the first place²¹. By defecting, they would only harm their own reputation, which puts them at risk to receive less cooperation in the future. Generous Scoring circumvents this risk by punishing defectors stochastically, with a well-defined probability. This probability is chosen such that the expected long-term loss in reputation exactly matches the short run gains from saving the cooperation costs. We note that this does not require the players to know all relevant game parameters in advance, or to explicitly calculate the respective probabilities. Instead, our simulations suggest that individuals may well be able to learn such strategies through elementary exploration and imitation processes.

Our results also suggest that direct and indirect reciprocity require different environments to emerge. Generous Tit-for-Tat requires players to interact sufficiently often, whereas Generous Scoring can also sustain cooperation when players only interact occasionally. However, for Generous Scoring to evolve, mutation rates need to be smaller than under direct reciprocity, and outside information needs to be sufficiently reliable (**Fig. 3 – Fig. 5**). While our results in the main text focus on simple first-order strategies, our general framework is equally applicable to more elaborate norms of indirect reciprocity. In particular, in **SI Section 6** we explore how our framework can be adapted to study strategies represented by finite-state automata with more than two states¹⁸ or the leading eight²², see also **Extended Data Fig. 6–Extended Data Fig. 9**.

We have explored how to make decisions when different sources of information are available. When individuals interact regularly, we find that they rely on direct information. They trust their own experiences more than indirect information which may be subject to noise. In contrast, when relationships are short-lived or superficial, cooperation can only be sustained when people act upon public reputations. Previous work suggests that indirect reciprocity requires social norms that are sufficiently complex^{21–23}. These norms make use of an unlimited regress; when assigning a new reputation to a person, observers need to take into account the reputation of the person’s co-player, which in turn depends on the reputation of the co-player’s previous interaction partner. Our model proposes a different view. To sustain cooperation, simple probabilistic rules based on a minimum of information can suffice.

Methods

In the following, we provide a more technical summary of our framework. We explain how it can be used to (i) derive the players' payoffs, (ii) characterise all Nash equilibria among reactive strategies, and (iii) study the co-evolution of direct and indirect reciprocity. For all details and proofs, we refer to the **SI**.

General framework. For the baseline model considered throughout the main text, we consider a game in a well-mixed population with n individuals. In each round, two individuals are randomly drawn to interact in one round of a prisoner's dilemma. They can either cooperate (C) or defect (D). Cooperation means to pay a cost $c > 0$ to provide a benefit $b > c$ to the co-player. Defection means to pay no cost, and for the co-player to gain no benefit. Both players decide independently. Their actions are observed by all population members. However, we assume indirect information is subject to perception errors: those members who only indirectly witness the interaction may misinterpret each player's action with probability ε . That is, with probability ε , outside observers take a C for a D , or vice versa. After the two interacting individuals have made their decisions, with probability d there is another round. In that case, again two individuals are chosen at random from the population to interact in a prisoner's dilemma. Otherwise, with probability $1 - d$, the game is over. The players' payoffs for the population game are defined as their average payoff over all rounds in which they participated in.

Each individual represents every other population member by a separate finite-state automaton. Each automaton can be in one of two possible states, 'good' (G) or 'bad' (B). The current state of the automaton depends on the individual's strategy, on the co-player's past actions, and on whether or not an error has occurred. In the main text, strategies are 4-tuples $(y, p, q, \lambda) \in [0, 1]$. The first entry y is the initial probability for the automaton to be in the good state. The second entry p and the third entry q are the conditional probabilities that the automaton is in the good state, given that the respective co-player just cooperated (defected) in a direct interaction, respectively. Finally, the value of λ is the probability that a player's indirect interactions with third parties are taken into account to update the automaton accordingly. For $\lambda = 0$, third-party interactions are completely ignored, and the automaton's state only depends on direct interactions. For $\lambda = 1$, every interaction of the respective co-player is equally taken into account, no matter whether or not the focal individual is directly involved. Individuals cooperate with those co-players they consider as good, and defect against those co-players they consider as bad.

We refer to the case of $\lambda = 0$ as direct reciprocity, and to $\lambda = 1$ as indirect reciprocity. We note that in exceptional cases, even a player with $\lambda = 1$ may base her decisions on direct experiences. This happens, for example, when the same two players are chosen to interact for two consecutive rounds. In that case, the players' second-round behavior will depend on their direct experience in the first round. In **SI Section 6.2**, we contrast this model with an alternative specification. In that alternative model, players who use indirect reciprocity ignore all direct information entirely. With minor modifications, all results presented herein carry over (see also **Extended Data Fig. 11**).

Derivation of a unified payoff equation of direct and indirect reciprocity. For our baseline framework, the players' payoffs can be calculated explicitly, without having to simulate the game. To derive the respective payoff equation, let each player i adopt some arbitrary but fixed strategy $(y_i, p_i, q_i, \lambda_i)$. Let $\bar{w} = 2/n$ denote the probability that a particular player is chosen to interact in the next round of the prisoner's dilemma. Similarly, $w = 2/(n(n-1))$ is the probability that a particular pair of players is chosen. Finally, we denote by $x_{ij}(t)$ the probability that player i considers player j to be good after t games have been played in the population. Given the value of $x_{ij}(t)$, we can

recursively compute $x_{ij}(t+1)$ as

$$\begin{aligned}
x_{ij}(t+1) &= (1-\bar{w}) x_{ij}(t) \\
&+ w (x_{ji}(t) p_i + (1-x_{ji}(t)) q_i) \\
&+ (\bar{w}-w) (1-\lambda_i) x_{ij}(t) \\
&+ w \lambda_i \sum_{l \neq i, j} \left((1-\varepsilon) x_{jl}(t) + \varepsilon (1-x_{jl}(t)) \right) p_i + \left((1-\varepsilon) (1-x_{jl}(t)) + \varepsilon x_{jl}(t) \right) q_i.
\end{aligned} \tag{6}$$

The summands on the right hand side reflect the following four possible events (illustrated in **Extended Data Fig. 1e–h**):

- (i) The first line on the right hand side corresponds to the case that player j does not interact at time t . This happens with probability $1 - \bar{w}$. In this case i 's automaton with respect to j does not change.
- (ii) The second line corresponds to the case that i directly interacts with j at time t . This happens with probability w . In this case, we assume players always get a chance to update the co-player's reputation. The updated reputation state depends on the values of p_i and q_i , and on j 's actions. Player j 's action is C with probability $x_{ji}(t)$ and D with probability $1-x_{ji}(t)$.
- (iii) The third line corresponds to the case that j interacts with some third party, which happens with cumulative probability $(\bar{w}-w)$, but player i decides not to react to this indirect information, with probability $(1-\lambda_i)$. In this case i 's automaton with respect to j does not change.
- (iv) The last line represents the case that j interacts with some third party l , which has probability w each, and i updates her automaton with respect to j accordingly. In this case, player i 's updated state depends on whether or not j cooperates, whether or not there is a perception error, and on the values of p_i and q_i . We sum up over all possible interactions of player j with third parties.

Given this recursion, we can calculate the value of $x_{ij}(t)$ for all future times t based on the initial condition $x_{ij}(0) = y_i$. This allows us to compute the weighted average $x_{ij} := (1-d) \sum_{t=0}^{\infty} d^t \cdot x_{ij}(t)$. This average corresponds to the probability to find player i 's automaton in the good state in a randomly picked round. Its value can be computed explicitly, by representing Eq. (6) in matrix notation (**SI Section 3**). Based on the values of x_{ij} , player i 's expected payoff becomes

$$\pi_i = \frac{1}{n-1} \sum_{j \neq i} (x_{ji} b - x_{ij} c). \tag{7}$$

This equation allows the explicit calculation of payoffs for arbitrary population compositions. Its results are in agreement with the payoffs that one obtains when simulating the game dynamics explicitly (**Extended Data Fig. 10**).

Equilibrium analysis. Based on the payoff formula (7), we can explicitly characterise the generic Nash equilibria of our model (i.e., those Nash equilibria that are robust with respect to small parameter changes). To this end, it is useful to introduce the variable δ , which is the pairwise continuation probability (the probability that two players interact again, given that they just had an interaction). This probability can be calculated explicitly. It depends on the population-wide continuation probability d and on the population size n , and it is given by $\delta = 2d / (2d + (n-1)n(1-d))$. For a derivation, see **SI Section 4**.

By extending the theory of zero-determinant strategies^{8–16}, we prove that a reactive strategy (y, p, q, λ) is a

generic Nash equilibrium for $0 < \delta < 1$ if it is either $ALLD = (0, 0, 0, \lambda)$, or if $p - q = r_\lambda^*$ with

$$r_\lambda^* := \frac{1 + (n-2)\delta\lambda}{1 + (n-2)(1-2\varepsilon)\lambda} \cdot \frac{c}{\lambda b}. \quad (8)$$

In **Fig. 3a,b**, the set of all strategies that satisfy $p - q = r_\lambda^*$ is depicted by a coloured face and a coloured dashed line, respectively. If the entire population adopts one of these Nash equilibrium strategies, no single player can gain a higher payoff by deviating.

We call a generic Nash equilibrium *cooperative*, if it has the additional property that all players are fully cooperative in the limit of rare errors. Due to this latter property, the strategy needs to satisfy

$$y = p = 1. \quad (9)$$

That is, the strategy always needs to assign a good reputation to unknown players, and to players who have cooperated in the latest relevant interaction. Combining Eqs. (8) and (9) shows that within the space of reactive strategies, there is exactly one cooperative Nash equilibrium of direct reciprocity ($\lambda = 0$). This strategy is *GTFT*, as defined in Eq. (1). Similarly, there is exactly one cooperative Nash equilibrium of indirect reciprocity ($\lambda = 1$), the strategy *GSCO*, defined by Eq. (2). In addition to these distinguished boundary cases, we can use Eqs. (8) and (9) to construct infinitely many cooperative Nash equilibria, one for every value of $\lambda \in [0, 1]$. We refer to the class of all these strategies as *Generous Reciprocators*. For all respective details, see **SI Section 4**.

Evolutionary analysis. We model the evolutionary spread of strategies in the population by a pairwise comparison process^{46,47}. Initially, players adopt an arbitrary strategy (y, p, q, λ) . Then one player is randomly chosen from the population to update her strategy. There are two distinct mechanisms how this updating can occur.

- (i) With probability μ , there is a mutation event. In that case, the focal player abandons her old strategy and instead switches to a new strategy (y', p', q', λ') . The first three entries, y', p', q' , are uniformly and independently drawn from the unit interval $[0, 1]$. For simplicity, we assume in most figures that the last entry λ' is either predetermined (for those simulations in which players are restricted to either direct or indirect reciprocity), or that it is randomly taken from the set $\{0, 1\}$.

In addition, in **Extended Data Fig. 12**, we explore how evolution operates when players can also adopt strategies with intermediate λ . For these simulations, we first compute how likely it is for a given strategy (y, p, q, λ) that a player's state with respect to a given co-player is updated between two consecutive games of the two players. The respective probability γ can be calculated as (**SI Section 5.4**)

$$\gamma = \frac{(n-2)\lambda}{1 + (n-2)\lambda}. \quad (10)$$

As one may expect, $\lambda = 0$ implies that $\gamma = 0$. That is, a player who ignores all third-party information only updates the co-player's state in a direct encounter and never in between. Similarly, $\lambda = 1$ implies that $\gamma = (n-2)/(n-1)$. That is, a player who takes all information into account has an $(n-2)/(n-1)$ chance to update the co-player's state before the two players interact again (the only exception occurs when the co-player engages in no third-party interaction in between, which happens with a $1/(n-1)$ probability). For the simulations shown in **Extended Data Fig. 12**, we randomly draw mutant strategies (y', p', q', λ') such that the respective γ' according to Eq. (10) is evenly distributed in $[0, (n-2)/(n-1)]$. In this way, we ensure that a randomly drawn mutant is approximately equally likely to base her decisions on direct and on third-party

information, respectively.

We note that alternatively, one could also consider a mutation scheme where λ itself is uniformly drawn from the unit interval $[0, 1]$. We do not employ this alternative mutation scheme here because the resulting mutant strategies would rarely engage in direct reciprocity. Intuitively, players in large populations have many more third-party interactions than they have direct interactions. As a result, even for a comparably small value of λ , the resulting γ according to Eq. (10) is typically close to one (especially if the population size n is large). For uniform λ , players would thus rarely act based on their direct experience with the respective co-player. For further details, see **SI Section 5.4**.

- (ii) With probability $1-\mu$, there is an imitation event. In that case, the focal player randomly chooses another player from the population as a potential role model. If the focal player's payoff according to Eq. (7) is given by π_F and the role model's payoff is π_R , the focal player adopts the role model's strategy with probability

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

The parameter $\beta \geq 0$ measures the strength of selection. For small values of β , the imitation probability is roughly $1/2$, independent of the strategies of the involved players. As the value of β increases, the more likely it becomes that the focal player only adopts those strategies that yield a higher payoff.

For positive values of μ and finite values of β , the two mechanisms of mutation and imitation give rise to an ergodic stochastic process on the space of all population compositions. To explore the evolutionary dynamics, we have simulated this process for a large number of updating events. We record which strategies the players adopt over time, and how often they cooperate. Because the process is ergodic, the time averages of these quantities converge, and they are independent of the initial population⁶⁷.

Specific methods employed for the figures. **Fig. 3a,b** depicts simulation results of the evolutionary process when all players are either required to use direct ($\lambda=0$) or indirect reciprocity ($\lambda=1$). We simulated the process for $2 \cdot 10^7$ mutant strategies. For panels **c-f**, we have looked at simulations where the initial population either employs noisy variant of *ALLD*, $(0.01, 0.01, 0.01, \lambda)$, or a conditionally cooperative strategy, $CC = (0.99, 0.99, 0.5, \lambda)$. We then recorded how long it takes on average until a mutant strategy reaches fixation, and which mutant strategy succeeds. Each bar depicts an average over 10^3 simulations. Panels **g,h** depict the players' payoffs when the population consists of a mixture of defectors and noisy discriminators (*TFT* in the case of direct reciprocity, *SCO* in the case with indirect reciprocity, with $p=0.99, q=0.01$). As parameters for this figure, we use $n=50, b=5, c=1, \varepsilon=0, \beta=10$, and the limit of rare mutations $\mu \rightarrow 0$.

Fig. 4 explores how different mutation rates affect the results of **Fig. 3**. As in **Fig. 3**, all players are either restricted to use direct ($\lambda=0$) or indirect reciprocity ($\lambda=1$). For the panels **a-f**, we have then simulated the evolutionary process for different continuation probabilities. Panels **a-c** depict which strategies the players use over time, for either the limit $\mu \rightarrow 0$ (grey bars) or a mutation rate of $\mu=0.01$ (coloured bars). The upper panels of **d-e** depict how much players cooperate on average, for different values of μ . The lower panels show how many different strategies are simultaneously present in a population on average. This number ranges from 1 in the limit of rare mutations to $n=50$ when mutations are abundant. Panels **g-i** consider populations that are initialised either with the same noisy variant of *ALLD* considered in **Fig. 3**, or with the same conditionally cooperative strategy. Again, we record how long it takes on average until the respective strategy has been removed from the population by the evolution-

ary process. Each data point represents an average of 50 independent simulations. Time is scaled such that units correspond to number of introduced mutant strategies. For $\mu \rightarrow 0$ these extinction times converge to the values in **Fig. 3c,d**. For $\mu \rightarrow 1$, extinction times converge to a value that is the same for both strategies and all continuation probabilities. Unless noted otherwise, parameters are the same as in **Fig. 3**.

Fig. 5 shows evolutionary results for the case that in addition to the y, p, q values, players are also free to choose $\lambda \in \{0, 1\}$. We consider two sets of simulations, either for the case of rare mutations ($\mu \rightarrow 0$), or for a positive mutation rate ($\mu = 0.02$). The panels **a–c** and **e–g** depict average trajectories for three specific scenarios. The scenarios differ in the game’s continuation probability and the error rate. The specific parameters we use are $\delta = 0.3, \varepsilon = 0.1$ (**Fig. 5a,e**), $\delta = 0.9, \varepsilon = 0.001$ (**Fig. 5b,f**), and $\delta = 0.999, \varepsilon = 0.01$ (**Fig. 5c,g**). The time trajectories represent an average over 1,000 simulations. In the initial population, all players employ $ALLD = (0, 0, 0, \lambda)$ with randomly chosen $\lambda \in \{0, 1\}$. Time is scaled such that units correspond to the number of introduced mutant strategies since the beginning of the simulation. In **Fig. 5d,h**, we systematically vary the continuation probability and the error rate of the game. Each data point corresponds to the time average of a single simulation with 10^7 time steps. Unless noted otherwise, parameters are the same as in **Fig. 3**.

Extended Data Fig. 3 explores how the results for the three scenarios considered in **Fig. 5** change as we vary four model parameters. The upper half of the figure considers the same basic setup as in **Fig. 5**: the population is initialised such that every player uses $ALLD$. Then we simulate the process for a sufficient time (at least until 10^7 mutant strategies have been introduced). This time is chosen such that the average cooperation rate and the average proportion of indirect reciprocity equilibrates, and that these quantities are independent of the chosen initial population. In the lower half we consider an alternative simulation scheme. Here, we take the average over 200 simulations with randomly chosen initial populations. Each simulation is only run for 10^5 time steps (mutant strategies introduced). The evolutionary parameters that we vary are the benefit-to-cost ratio b/c (between 1 and 12), population size n (between 2 and 1,024), selection strength β (between 0.01 and 100) and mutation rate μ (between 0.0001 and 1). As the baseline parameters, we use the same values as in **Fig. 5a–c**.

Extended Data Fig. 4 investigates in more detail the non-monotonicity of evolving cooperation rates in **Extended Data Fig. 3e**. To this end, we again consider the scenario with intermediately many interactions and reliable information ($\delta = 0.9, \varepsilon = 0.001$, orange curve in **Extended Data Fig. 3e**). Panels **a,b**, explore how many mutant strategies it takes on average to invade two different resident strategies. Because the non-monotonicity arises in a scenario that favours the evolution of indirect reciprocity, we consider two residents with $\lambda = 1$. The defector resident is given by $(0.001, 0.001, 0.001, 1)$, whereas the cooperative resident adopts the strategy $(0.999, 0.999, 0.650, 1)$. We have run 1,000 simulations for different values of the selection strength parameter ($10^{-3} \leq \beta \leq 10^2$). Dots represent outcomes of individual simulations, whereas the red and blue curve represent average values. For panels **c–e**, we have run the same evolutionary process as in **Extended Data Fig. 3e** for the scenario with intermediately many interactions and reliable information. For three different intensities of selection, we have recorded the distribution of cooperation rates over a simulation with $2 \cdot 10^7$ time steps. Time steps are measured in number of mutant strategies introduced by the process.

For **Extended Data Fig. 5**, we have re-run the simulations in **Fig. 5d** for different noise scenarios. Except for the changes explicitly mentioned (by changing the error scenario, or the information available to the players), the simulations have been performed exactly as for **Fig. 5d**.

Extended Data Fig. 6 explores the stability of three different finite-state automata against a single *ALLD* or *ALLC* mutant. Because there is no efficient payoff formula that computes the payoffs of arbitrary automata in the context of indirect reciprocity with noisy observations, we computed the payoffs by simulations. To this end, we assumed all automata are initialized in the good state. Then players engage in $2 \cdot 10^6$ pairwise interactions. To compute the players' average payoffs, we take the mean over all their payoffs in the second half of the simulation, as in previous work³². Taking the average over all rounds would not alter our conclusions. As game parameters, we use $n = 50$, $b = 5$, $c = 1$, and $\varepsilon = 0.05$.

For **Extended Data Fig. 7**, we have first simulated the players' payoffs for all possible population compositions (k_A, k_C, k_D) . Here, k_A is the number of players who adopt the respective automaton strategy, k_C is the number of unconditional cooperators, and k_D is the number of defectors. For these payoff calculations, we have employed the same process as in **Extended Data Fig. 6**. For pre-computed payoffs, the fixation probability of a given mutant strategy into any other resident strategy can be computed explicitly⁶⁸. Based on all pairwise fixation probabilities, one can then compute how often each strategy is played on average⁴⁸. This yields the panels **a–c**. For positive mutation rates, the abundance of each strategy can still be computed explicitly, by formulating the evolutionary process as a Markov chain. The states of this Markov chain are all possible population compositions (k_A, k_C, k_D) . When $n = 50$, there are 1,326 such states. Thus, the dynamics can be described by an $1,326 \times 1,326$ transition matrix. The entries of this transition matrix describe with which probability the population moves from state (k_A, k_C, k_D) to state (k'_A, k'_C, k'_D) after one evolutionary updating event (see for example SI Section 6.2 in Ref. 32). The invariant distribution of this Markov chain can be computed directly. It describes how often each state is visited by the evolutionary process, as illustrated in panels **d–f**. Based on this invariant distribution, we can also calculate how often players cooperate on average, as shown in panels **g–i**.

Extended Data Fig. 8 and **Extended Data Fig. 9** use exactly the same method as **Extended Data Fig. 6** and **Extended Data Fig. 7**, respectively. As the only difference, the finite-state automata are replaced by leading-eight strategies.

Extended Data Fig. 10 considers a population consisting of 49 conditional cooperators and a single defector. The cooperators employ the strategy $(1, 1, 0.01, \lambda)$, whereas the remaining defector applies the strategy $(0, 0, 0, \lambda)$. We use two independent approaches to compute the players' payoffs, the payoff equation (7), and explicit simulations of the game dynamics (we averaged over 10^5 iterations per parameter combination). The parameters are **a**, $\varepsilon = 0.001$, $\delta = 0.9$ and **b**, $\varepsilon = 0.45$, $\delta = 0.999$. The respective Python scripts used to run the simulations and for solving Eq. (7) are provided in the SI.

For **Extended Data Fig. 11**, we repeat the simulations done in **Fig. 5a–d**, but now using the alternative strategy set (y, p, q, κ) described in **SI Section 6.2**. Here, players who employ indirect reciprocity ignore all direct information they may have. Game parameters are the same as in **Fig. 5a–d**.

Finally, in **Extended Data Fig. 12**, we repeat the simulations in **Fig. 3**, **Fig. 4**, and **Fig. 5**, but now allowing for intermediate values of λ . In **Extended Data Fig. 12a–d**, we consider the case of a fixed λ value. To this end, we use five different values of λ , which according to Eq. (10) map to the values of $\gamma \in \{0, 1/4, 1/2, 3/4, \gamma_{\max}\}$, with $\gamma_{\max} = (n-2)/(n-1)$. For the panels **a**, **b**, we consider the same setup as in **Fig. 3a,b**, and use the same parameter

values. For the panels **c,d**, we use the same setup and the same parameters as in **Fig. 4d,f**. Finally, for the panels **e–h**, we use the same parameters and the same general setup as in **Fig. 5a–d**. However, while in **Fig. 5a–d**, players are restricted to strategies with either $\lambda = 0$ or $\lambda = 1$, here they can adopt arbitrary strategies with $0 < \lambda < 1$. The bottom panels of **Extended Data Fig. 12e–g** show how often residents adopt different values of γ by the end of each simulation (for 1,000 simulations in total). For details on how the respective mutant strategies are generated, see **Methods** and **SI Section 5.4**.

Data availability. The raw data generated for the main text, which was used to create **Fig. 3–Fig. 5**, is available at https://osf.io/brnvx/?view_only=4adc0b791a3640df88c94362d0f164e6. The raw data for the Extended Data Figures is available from the authors upon request.

Code availability. All simulations and numerical calculations have been performed with MATLAB R2014A and Python 2.7. In the **SI Appendix**, we provide the Python scripts used to simulate the game dynamics, numerically calculate the players’ expected payoffs, and simulate the evolutionary process.

References

- [1] Trivers, R. L. The evolution of reciprocal altruism. *The Quarterly Review of Biology* **46**, 35–57 (1971).
- [2] Sugden, R. *The economics of rights, co-operation and welfare* (Blackwell, Oxford, UK, 1986).
- [3] Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
- [4] Sigmund, K. *The Calculus of Selfishness* (Princeton Univ. Press, Princeton, NJ, 2010).
- [5] Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
- [6] Nowak, M. A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
- [7] Hauert, C. & Schuster, H. G. 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 (1997).
- [8] Press, W. H. & Dyson, F. D. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. *PNAS* **109**, 10409–10413 (2012).
- [9] 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).
- [10] Stewart, A. J. & Plotkin, J. B. Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences USA* **111**, 17558 – 17563 (2014).
- [11] Szolnoki, A. & Perc, M. Evolution of extortion in structured populations. *Physical Review E* **89**, 022804 (2014).
- [12] 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).
- [13] Pan, L., Hao, D., Rong, Z. & Zhou, T. Zero-determinant strategies in iterated public goods game. *Scientific Reports* **5**, 13096 (2015).
- [14] Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: Zero-determinant strategies in noisy games. *Physical Review E* **91**, 052803 (2015).
- [15] McAvoy, A. & Hauert, C. Autocratic strategies for iterated games with arbitrary action spaces. *Proceedings of the National Academy of Sciences* **113**, 3573–3578 (2016).

- [16] Ichinose, G. & Masuda, N. Zero-determinant strategies in finitely repeated games. *Journal of Theoretical Biology* **438**, 61–77 (2018).
- [17] Hilbe, C., Chatterjee, K. & Nowak, M. A. Partners and rivals in direct reciprocity. *Nature Human Behaviour* **2**, 469–477 (2018).
- [18] 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).
- [19] Reiter, J. G., Hilbe, C., Rand, D. G., Chatterjee, K. & Nowak, M. A. Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nature Communications* **9**, 555 (2018).
- [20] Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998).
- [21] Leimar, O. & Hammerstein, P. Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society B* **268**, 745–753 (2001).
- [22] Ohtsuki, H. & Iwasa, Y. How should we define goodness? – Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology* **231**, 107–20 (2004).
- [23] Santos, F. P., Santos, F. C. & Pacheco, J. M. Social norm complexity and past reputations in the evolution of cooperation. *Nature* **555**, 242–245 (2018).
- [24] Sigmund, K. Moral assessment in indirect reciprocity. *Journal of Theoretical Biology* **299**, 25–30 (2012).
- [25] Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters* **71**, 397–404 (2001).
- [26] Grujic, J. *et al.* A comparative analysis of spatial prisoner’s dilemma experiments: Conditional cooperation and payoff irrelevance. *Scientific Reports* **4**, 4615 (2014).
- [27] Wedekind, C. & Milinski, M. Cooperation through image scoring in humans. *Science* **288**, 850–852 (2000).
- [28] Okada, I., Yamamoto, H., Sato, Y., Uchida, S. & Sasaki, T. Experimental evidence of selective inattention in reputation-based cooperation. *Scientific Reports* **8**, 14813 (2018).
- [29] Molleman, L., van den Broek, E. & Egas, M. Personal experience and reputation interact in human decisions to help reciprocally. *Proceedings of the Royal Society B* **28**, 20123044 (2013).
- [30] 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).
- [31] Uchida, S. & Sasaki, T. Effect of assessment error and private information on stern-judging in indirect reciprocity. *Chaos, Solitons and Fractals* **56**, 175–180 (2013).
- [32] Hilbe, C., Schmid, L., Tkadlec, J., Chatterjee, K. & Nowak, M. A. Indirect reciprocity with private, noisy, and incomplete information. *Proceedings of the National Academy of Sciences USA* **115**, 12241–12246 (2018).
- [33] Raub, W. & Weesie, J. Reputation and efficiency in social interactions: An example of network effects. *American Journal of Sociology* **96**, 626–654 (1990).
- [34] Pollock, G. & Dugatkin, L. A. Reciprocity and the emergence of reputation. *Journal of Theoretical Biology* **159**, 25–37 (1992).
- [35] Roberts, G. Evolution of direct and indirect reciprocity. *Proceedings of the Royal Society B* **275**, 173–179 (2007).
- [36] Nakamaru, M. & Kawata, M. Evolution of rumours that discriminate lying defectors. *Evolutionary Ecology Research* **6**, 261–283 (2004).
- [37] Seki, M. & Nakamaru, M. A model for gossip-mediated evolution of altruism with various types of false information by speakers and assessment by listeners. *Journal of Theoretical Biology* **407**, 90–105 (2016).
- [38] Ohtsuki, H. Reactive strategies in indirect reciprocity. *Journal of Theoretical Biology* **227**, 299–314 (2004).

- [39] Nowak, M. A. & Sigmund, K. The dynamics of indirect reciprocity. *Journal of Theoretical Biology* **194**, 561–574 (1998).
- [40] Berger, U. Learning to cooperate via indirect reciprocity. *Games and Economic Behavior* **72**, 30–37 (2011).
- [41] Brandt, H. & Sigmund, K. The logic of reprobation: Assessment and action rules for indirect reciprocation. *Journal of Theoretical Biology* **231**, 475–486 (2004).
- [42] Uchida, S. Effect of private information on indirect reciprocity. *Physical Review E* **82**, 036111 (2010).
- [43] Martinez-Vaquero, L. A. & Cuesta, J. A. Evolutionary stability and resistance to cheating in an indirect reciprocity model based on reputation. *Physical Review E* **87**, 052810 (2013).
- [44] Nakamura, M. & Masuda, N. Indirect reciprocity under incomplete observation. *PLoS Computational Biology* **7**, e1002113 (2011).
- [45] Tanabe, S., Suzuki, H. & Masuda, N. Indirect reciprocity with trinary reputations. *Journal of Theoretical Biology* **317**, 338–347 (2013).
- [46] Szabó, G. & Tóke, C. Evolutionary Prisoner’s Dilemma game on a square lattice. *Physical Review E* **58**, 69–73 (1998).
- [47] 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).
- [48] Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory* **131**, 251–262 (2006).
- [49] Fudenberg, D., Nowak, M. A., Taylor, C. & Imhof, L. A. Evolutionary game dynamics in finite populations with strong selection and weak mutation. *Theoretical Population Biology* **70**, 352–363 (2006).
- [50] Imhof, L. A. & Nowak, M. A. Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B* **277**, 463–468 (2010).
- [51] Wu, B., Gokhale, C. S., Wang, L. & Traulsen, A. How small are small mutation rates? *Journal of Mathematical Biology* **64**, 803–827 (2012).
- [52] McAvoy, A. Comment on “Imitation processes with small mutations”. *J. Econ. Theory* **159**, 66–69 (2015).
- [53] Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences USA* **102**, 10797–10800 (2005).
- [54] García, J. & Traulsen, A. The structure of mutations and the evolution of cooperation. *PLoS One* **7**, e35287 (2012).
- [55] van Segbroeck, S., Pacheco, J. M., Lenaerts, T. & Santos, F. C. Emergence of fairness in repeated group interactions. *Physical Review Letters* **108**, 158104 (2012).
- [56] 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).
- [57] Stewart, A. J. & Plotkin, J. B. The evolvability of cooperation under local and non-local mutations. *Games* **6**, 231–250 (2015).
- [58] Santos, F. P., Santos, F. C. & Pacheco, J. M. Social norms of cooperation in small-scale societies. *PLoS Computational Biology* **12**, e1004709 (2016).
- [59] Hauser, O., Hilbe, C., Chatterjee, K. & Nowak, M. A. Social dilemmas among unequals. *Nature* **572**, 524–527 (2019).
- [60] Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907 (2007).
- [61] Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863 (2010).

- [62] García, J. & Traulsen, A. Leaving the loners alone: Evolution of cooperation in the presence of antisocial punishment. *Journal of Theoretical Biology* **307**, 168–173 (2012).
- [63] Hauert, C. & Imhof, L. Evolutionary games in deme structured, finite populations. *Journal of Theoretical Biology* **299**, 106–112 (2012).
- [64] Lee, Y., Iwasa, Y., Dieckmann, U. & Sigmund, K. Social evolution leads to persistent corruption. *Proceedings of the National Academy of Sciences USA* **116**, 13276–13281 (2019).
- [65] Brandt, H. & Sigmund, K. The good, the bad and the discriminator - errors in direct and indirect reciprocity. *Journal of Theoretical Biology* **239**, 183–194 (2006).
- [66] Selten, R. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* **4**, 25–55 (1975).
- [67] Karlin, S. & Taylor, H. M. A. *A First Course in Stochastic Processes* (Academic, London, 1975), 2nd edition edn.
- [68] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

Acknowledgments.

This work was supported by the European Research Council Start Grant 279307: Graph Games (to K.C.), the Austrian Science Fund (FWF) NFN Grant S11407-N23 Rigorous Systems Engineering/Systematic Methods in Systems Engineering (to K.C.), and the European Research Council Starting Grant 850529: E-DIRECT (to C.H.).

Author contributions.

All authors conceived the study, performed the analysis, discussed the results and wrote the manuscript.

Competing interests.

The authors declare no competing interests.

Additional information.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to L.S.

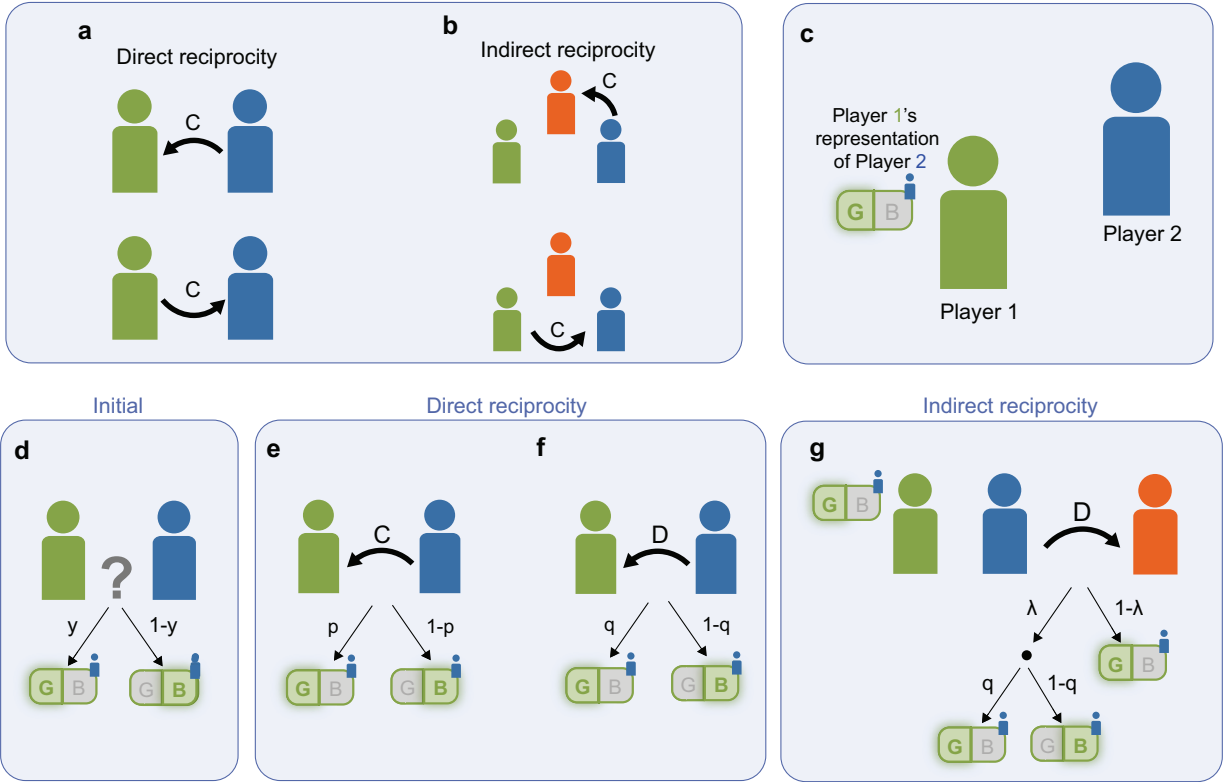


Figure 1: A unifying framework for direct and indirect reciprocity. **a**, Under direct reciprocity, an individual's cooperation is returned directly by the beneficiary. **b**, Under indirect reciprocity, cooperation is not returned by the beneficiary, but by some observer. **c**, To model direct and indirect reciprocity we consider individuals who assign one of two possible reputations to their co-player, good (*G*) or bad (*B*). The current assignment is highlighted in green. Individuals cooperate (*C*) with those co-players they consider as good, and they defect (*D*) against those they deem as bad. **d–g**, Whether an individual considers a co-player as good depends on her strategy (y, p, q, λ). Here, y is the initial probability to assign a good reputation to the co-player, without having any information; p and q are the probabilities to assign a good reputation after the co-player has cooperated or defected in a direct interaction, respectively. The receptivity λ is the probability with which an individual takes third-party interactions of the respective co-player into account. For $\lambda = 0$, we obtain a model of direct reciprocity. For $\lambda = 1$, we obtain a model of indirect reciprocity. While the illustrations depict one-way interactions for simplicity, our model considers two-way interactions. When two players are chosen to interact, they both decide simultaneously whether to cooperate or defect. All other population members observe their choices.

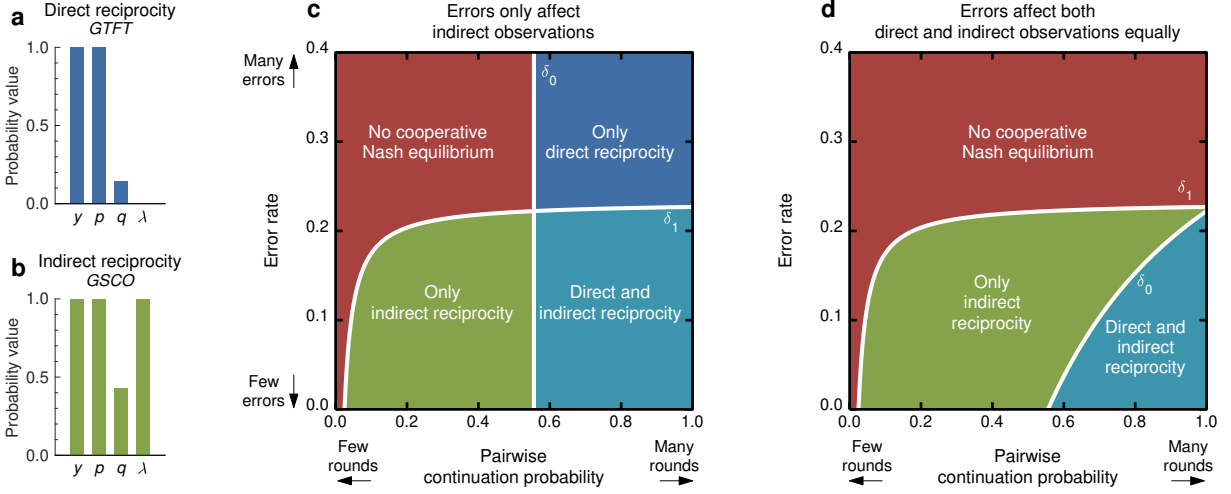


Figure 2: An equilibrium analysis reveals when direct or indirect reciprocity can sustain cooperation. **a,b**, Within the reactive strategies, there is one cooperative Nash equilibrium for direct reciprocity (*GTFT*), and one such equilibrium for indirect reciprocity (*GSCO*). Both strategies have in common that they always cooperate in the first round, or if the co-player has cooperated in the last relevant interaction ($y = p = 1$). They differ in how they react to a co-player's defection, as described by Eqs. (1) and (2), and in whether they take into account indirect information. **c**, Depending on the parameters of the game, there are up to four scenarios: (i) When there are few rounds and many perception errors, cooperation is infeasible; (ii) When there are intermediately many rounds and few perception errors, cooperation can be sustained by indirect but not by direct reciprocity; (iii) When there are many rounds and many perception errors, cooperation can be sustained by direct but not by indirect reciprocity; (iv) When there are many rounds and few errors, both direct and indirect reciprocity support cooperation. **d**, In case direct and indirect observations are subject to the same error rate, there is no region in which direct reciprocity can sustain cooperation but indirect reciprocity cannot. The figure shows the case of $n = 50$, $b = 1.8$ and $c = 1$. In **c**, the white lines depict the continuation probabilities δ_0 and δ_1 given by Eq. (3). In **d**, they are given by $\delta_0 = c / ((1 - 2\varepsilon)b)$ and $\delta_1 = c / ((n - 1)(1 - 2\varepsilon)b - (n - 2)c)$, where ε is now the joint error probability for both direct and indirect observations.

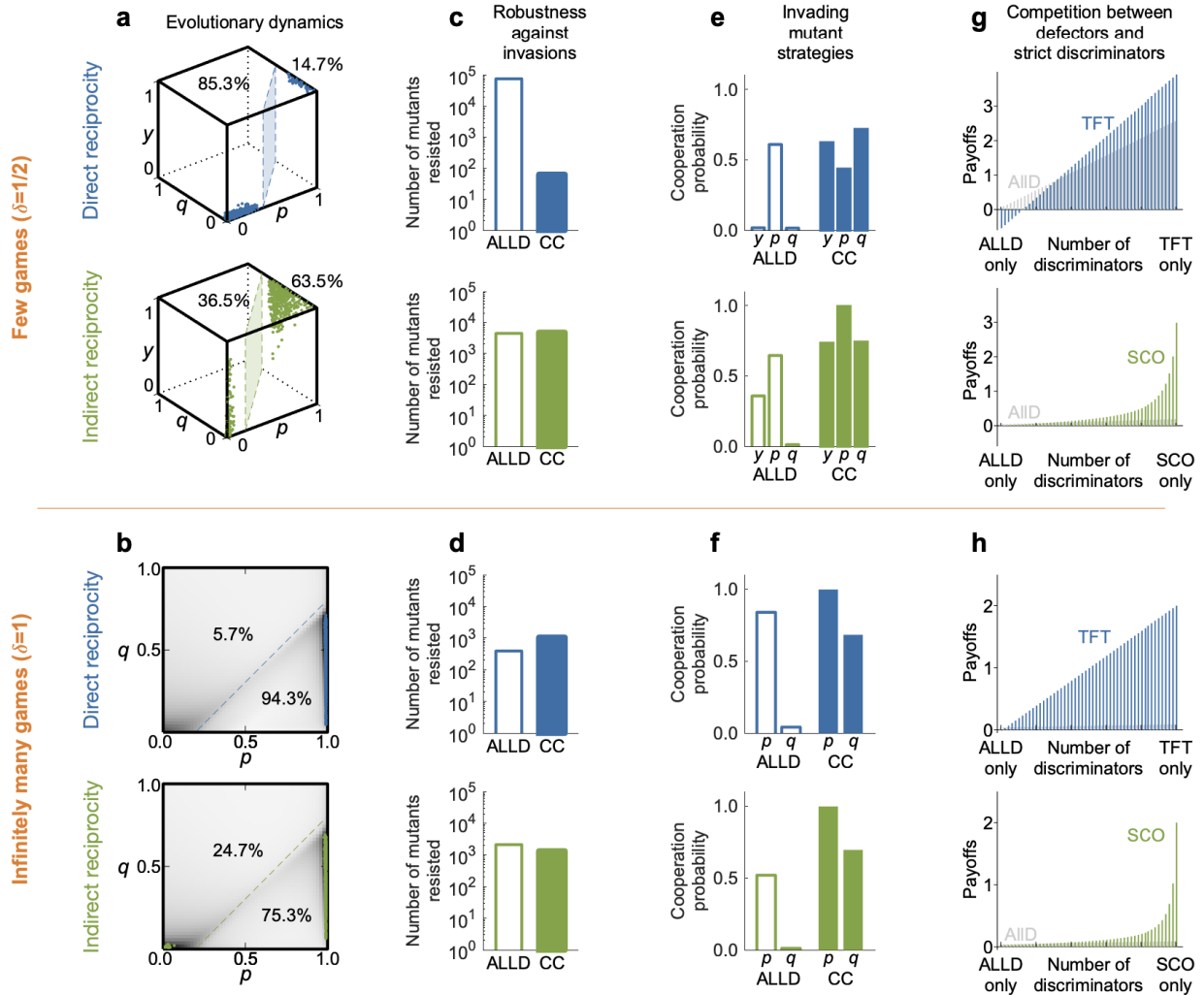


Figure 3: Evolutionary dynamics of direct and indirect reciprocity. We use individual based simulations to explore the dynamics if either all players engage in direct (blue) or indirect reciprocity (green). We consider two scenarios: individuals either engage in only a few games (top) or in infinitely many games (bottom). **a,b**, We find that over the course of evolution, populations cluster in two regions of the strategy space. Populations are either in the vicinity of *ALLD* (where $y \approx p \approx q \approx 0$), or in the vicinity of conditionally cooperative strategies (where $p \approx 1$). Percentages represent the fraction of time spent in each of these two neighbourhoods. Dots represent the 500 most long-lived resident strategies. As the impact of the first round is negligible for $\delta = 1$, the state space degenerates to a square instead of a cube. **c,d**, We have recorded how many mutants it takes to invade a population of defectors or conditional cooperators. We find that a larger number of rounds undermines the stability of *ALLD*, and enhances the stability of the cooperators. **e,f**, In addition, we have recorded which mutant strategies invade these two resident strategies. On average, defectors are invaded by conditionally cooperative strategies with $p \gg q$. **g,h** Under direct reciprocity, the payoff of a discriminating mutant (*TFT*) in an *ALLD* population increases linearly in the number of mutants. Under indirect reciprocity, the payoff of a discriminating mutant (*SCO*) is nonlinear. As baseline parameters in our evolutionary simulations, we use $n = 50$ and $b/c = 1$. For the exact setup of these simulations, see **Methods**.

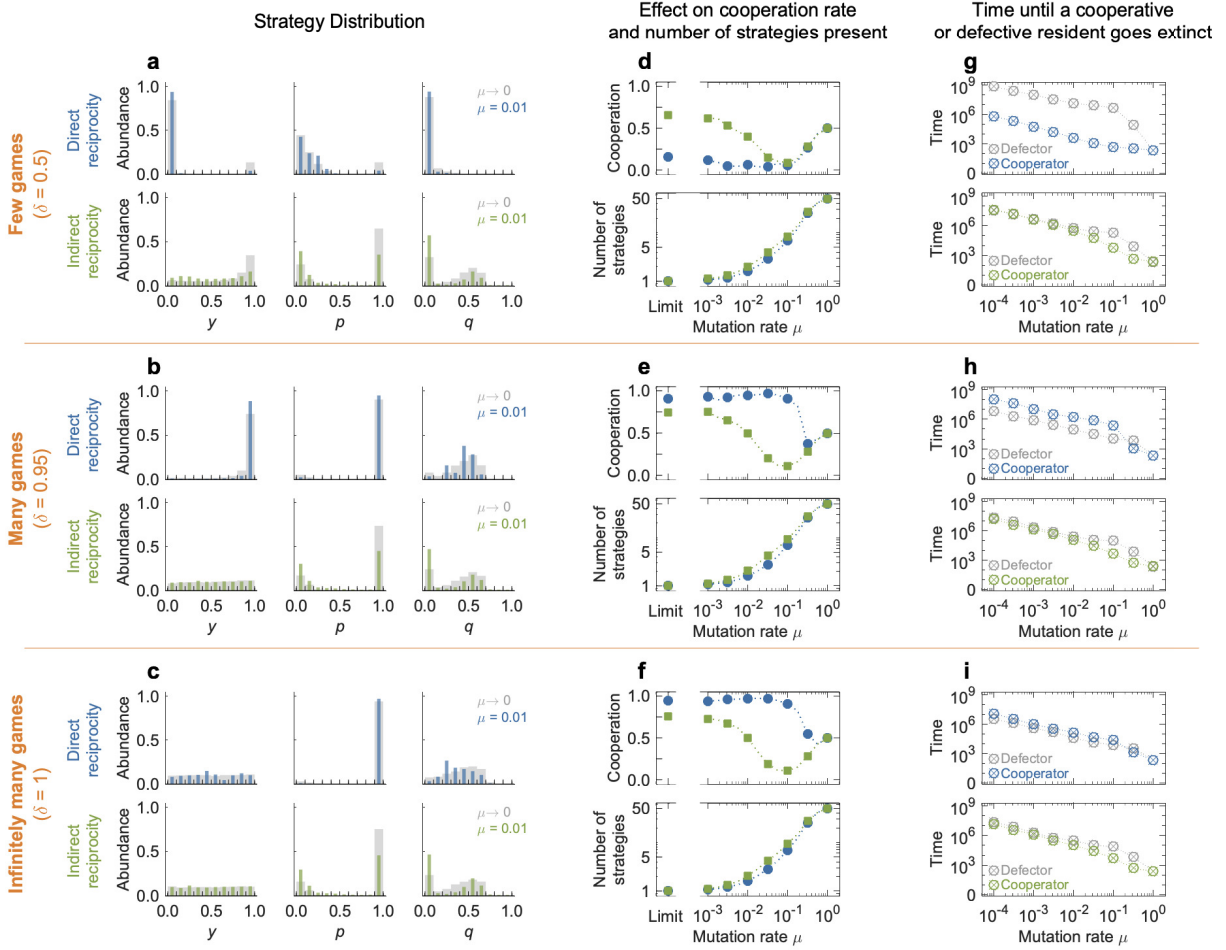


Figure 4: Impact of mutations on either direct or indirect reciprocity. We have run additional simulations to explore how larger mutation rates affect the results in Fig. 3. We consider the same two scenarios with few games and infinitely many games, and in addition a scenario where the number of games is large but finite. **a–c**, We have first run simulations for a particular positive mutation rate (coloured). We compare them with the results for the limit of rare mutations (grey). The bar diagrams depict how often we are to observe players to use strategies (y , p , q , λ) either for direct or indirect reciprocity. Similar to Fig. 3, we find that players are clustered in two regions of the strategy space. Either they tend to defect ($y \approx q \approx 0$) or they are conditionally cooperative ($p \approx 1$, $q < 1$). We note that the scenario for infinitely many games yields similar results as the scenario with many games. However, now the initial propensity to cooperate y becomes irrelevant. **d–f**, When we vary the mutation rate systematically, we find that while cooperation is relatively stable for direct reciprocity, cooperation under indirect reciprocity is reduced (upper panels). Interestingly, the number of different strategies that are simultaneously present in the population only differ marginally between direct and indirect reciprocity (lower panels). **g–i**, To explore what would cause this reduction in cooperation for indirect reciprocity, we have checked the stability of defectors and conditional cooperators for various mutation rates, as in Fig. 3c,d. In the interval $0.01 < \mu < 0.1$ where indirect reciprocity yields the lowest cooperation rate, we find that the stability of defectors is enhanced, whereas the stability of cooperators is reduced. For the exact setup of these simulations, see **Methods**.

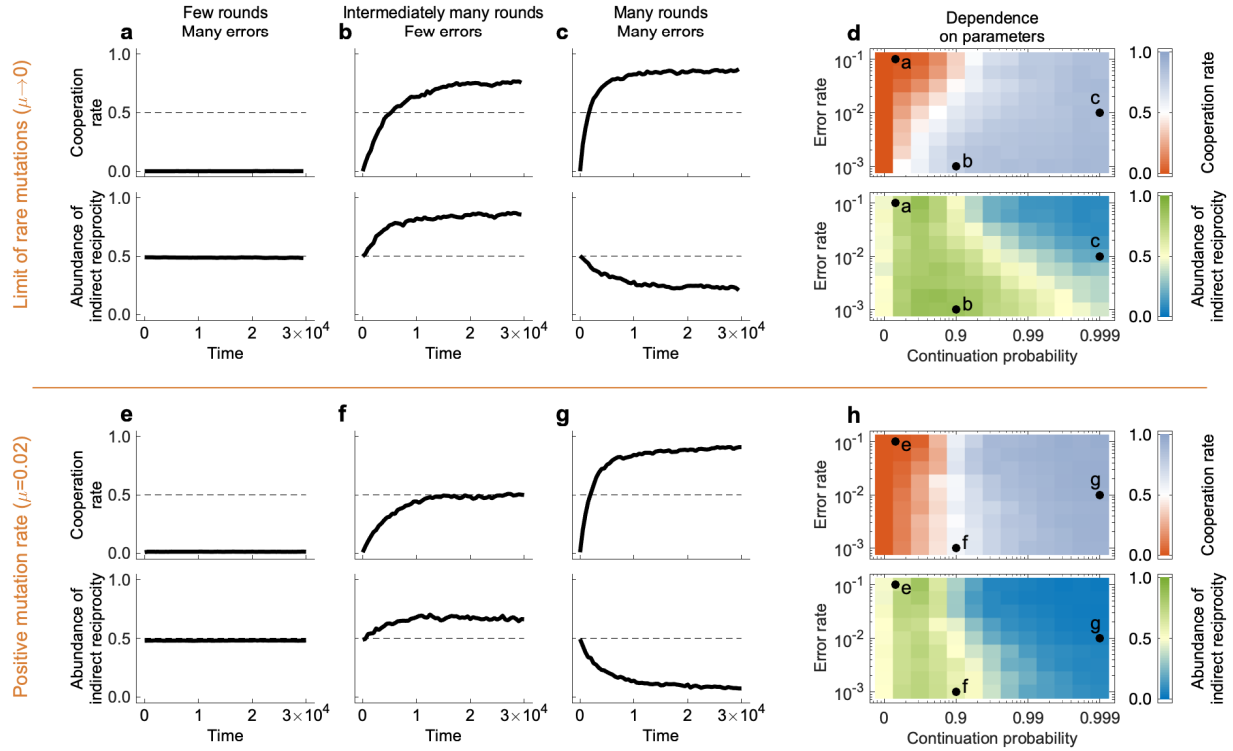
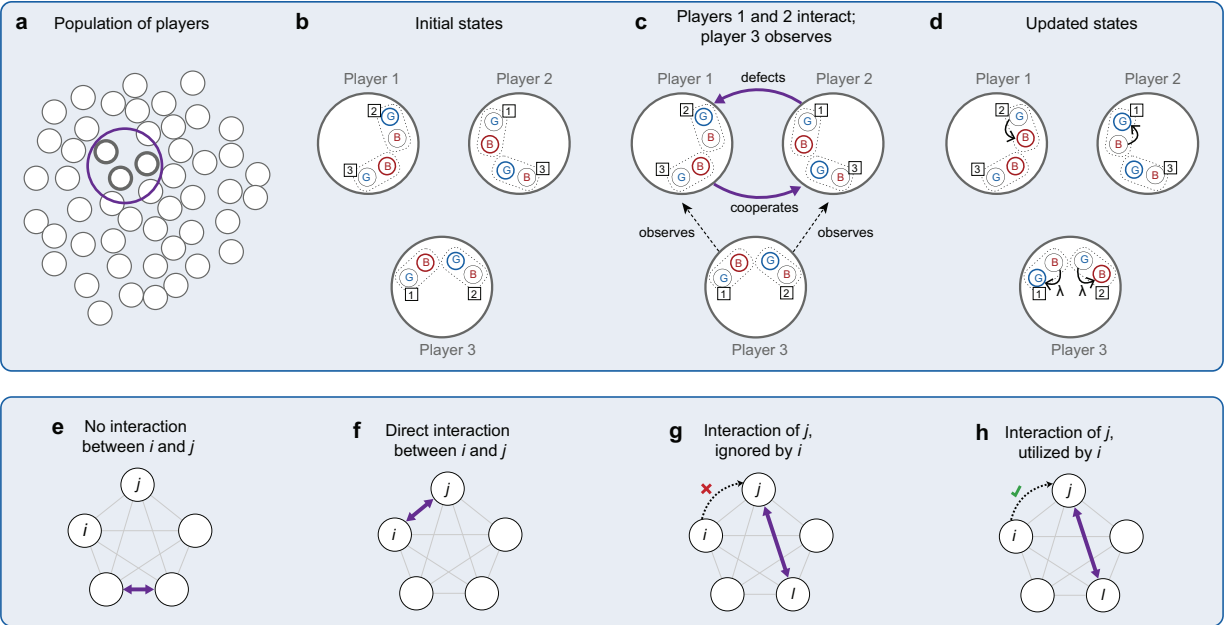
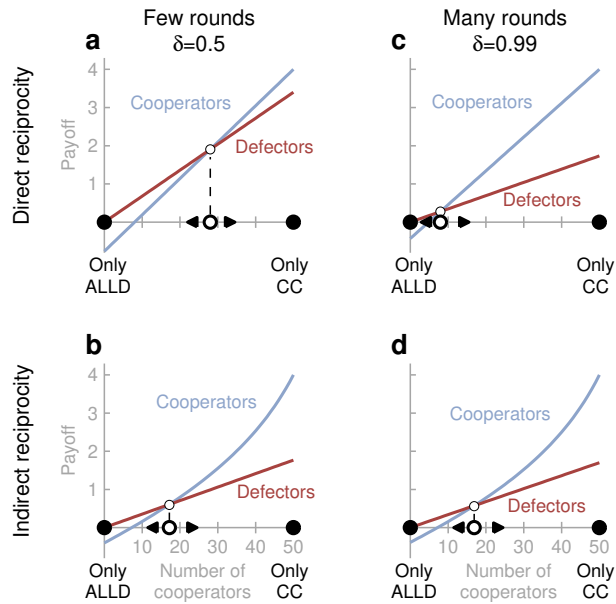


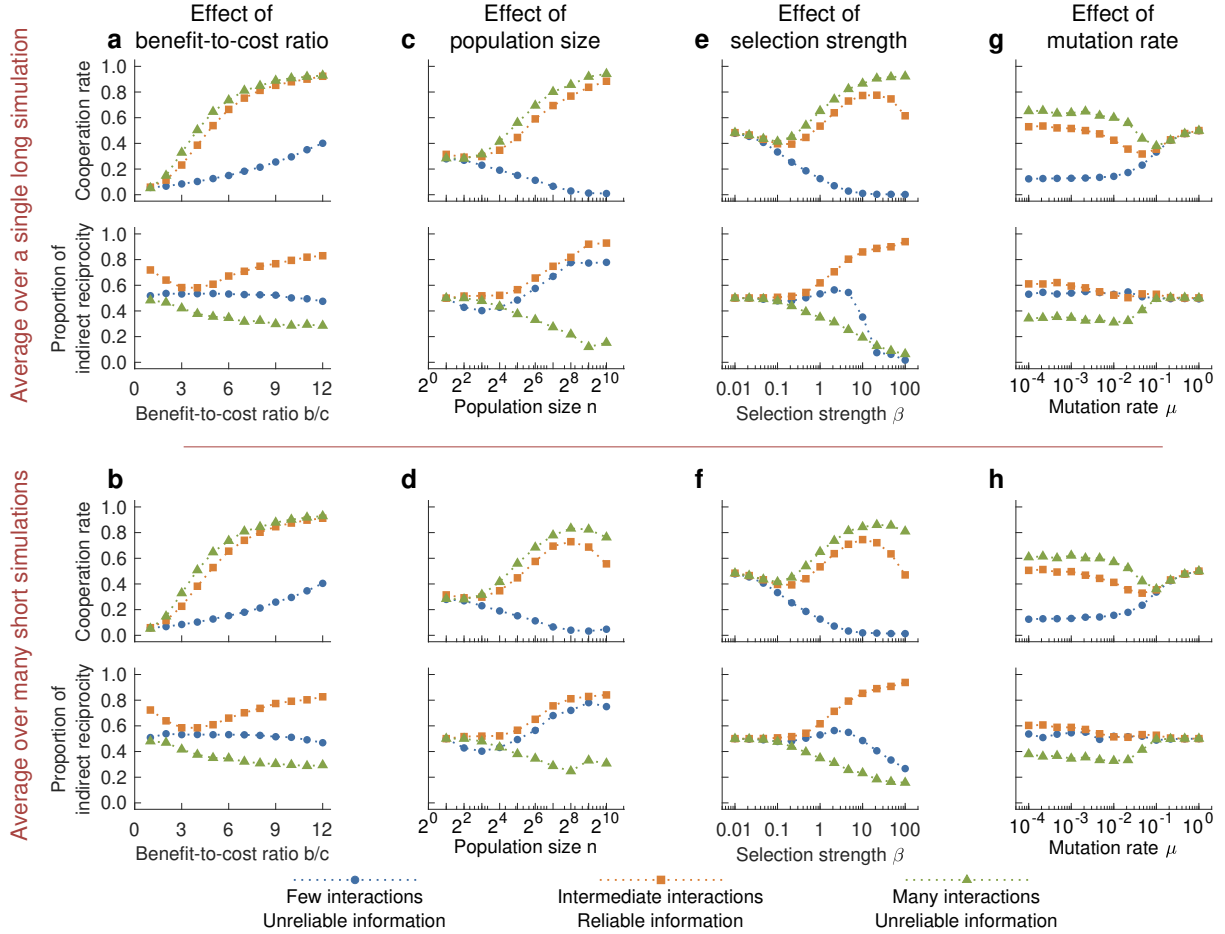
Figure 5: The coevolution of conditional cooperation and information use. To explore when individuals themselves learn to use indirect information, we ran simulations in which players can either use direct information only ($\lambda = 0$) or all information ($\lambda = 1$). **a–c**, We started with three particular scenarios in the limit of rare mutations. The scenarios differ in how often subjects interact on average and how noisy indirect information is. When there are only a few interactions and considerable noise, cooperation does not evolve at all (**a**). In the other two scenarios, cooperation either evolves due to indirect (**b**) or to direct (**c**) reciprocity. **d**, In a next step, we systematically varied the continuation probability δ and the error rate ε . Again, indirect reciprocity evolves for intermediate continuation probabilities. **e–g**, We obtain qualitatively similar results for positive mutation rates ($\mu = 0.02$). **h**, However, the green region in which individuals take into account indirect information is substantially diminished. For the exact setup of these simulations, see **Methods**.



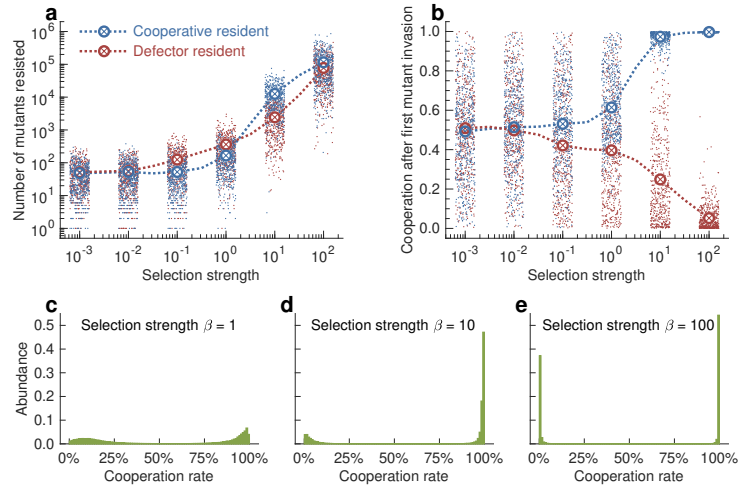
Extended Data Figure 1: Schematic representation of the model. **a**, We consider a population of size n . To illustrate the basic workings of our model, we focus on three arbitrary players. **b**, Each player has a separate finite-state automaton with two possible states G and B for each co-player. The current state is marked in bold. In this example, player 1 considers player 2 as good and player 3 as bad. **c**, In each round, two players are chosen at random to interact in a prisoner's dilemma. Players cooperate if they consider their co-player to be good and they defect otherwise. The other population members do not participate in the game, but they observe its outcome. **d**, After the interaction, both active players update their respective automata, depending on their strategy and on the co-player's action. In addition, each observer independently updates her automata with respect to players 1 and 2 with probability λ each. **e-h**, We can mathematically describe how player i 's automaton with respect to player j changes over time by distinguishing four possible events. First, player j is not chosen to interact, such that player i 's automaton remains unaffected (**e**); second, players i and j interact with each other and update their respective states accordingly (**f**); third, player j interacts with someone else, but player i does not take this interaction into account (**g**); fourth, player j interacts with someone else, and player i updates j 's state accordingly (**h**).



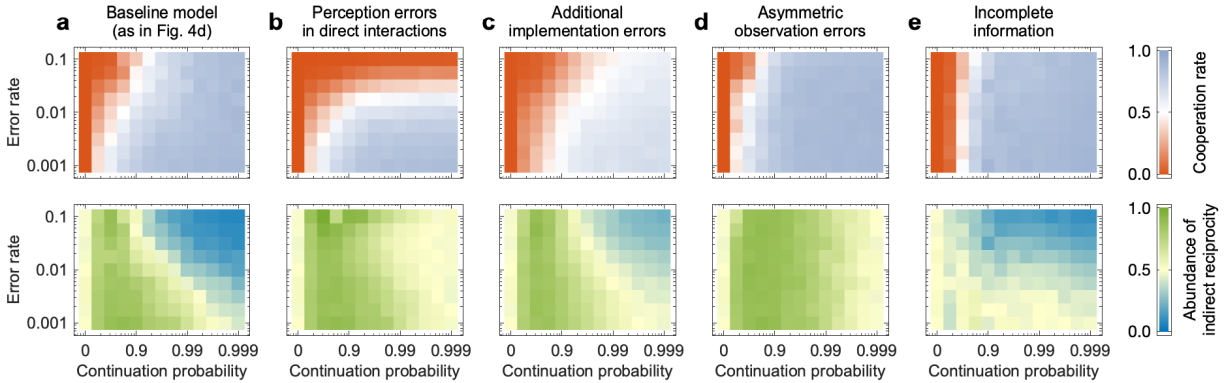
Extended Data Figure 2: Competition between conditional cooperators and defectors. We compare the performance of conditional cooperators with strategy $(1, 1, 1/3, \lambda)$ in a population of defectors, $(0, 0, 0, \lambda)$. We consider four scenarios, depending on whether players use direct or indirect reciprocity and depending on whether pairs interact only a few times or often. Each panel shows the payoff of cooperators and defectors depending on how many of the 50 population members are cooperators, for $b = 5$ and $c = 1$. In all four cases we find bistability (as indicated by the arrows on the x -axis). That is, defectors have the higher payoff when there are few cooperators and the lower payoff when there are many cooperators. However, the threshold number of cooperators necessary to make cooperation beneficial differs. Indirect reciprocity has the lower threshold when there are only few rounds, because cooperators are better able to restrict the payoff of defectors (as indicated by the smaller slope of the red line in **b** compared to **a**). Direct reciprocity has the lower threshold when there are many rounds. Here, already a few cooperators suffice to invade the defectors. In contrast, for indirect reciprocity cooperators need to establish a critical mass because their payoffs increase nonlinearly.



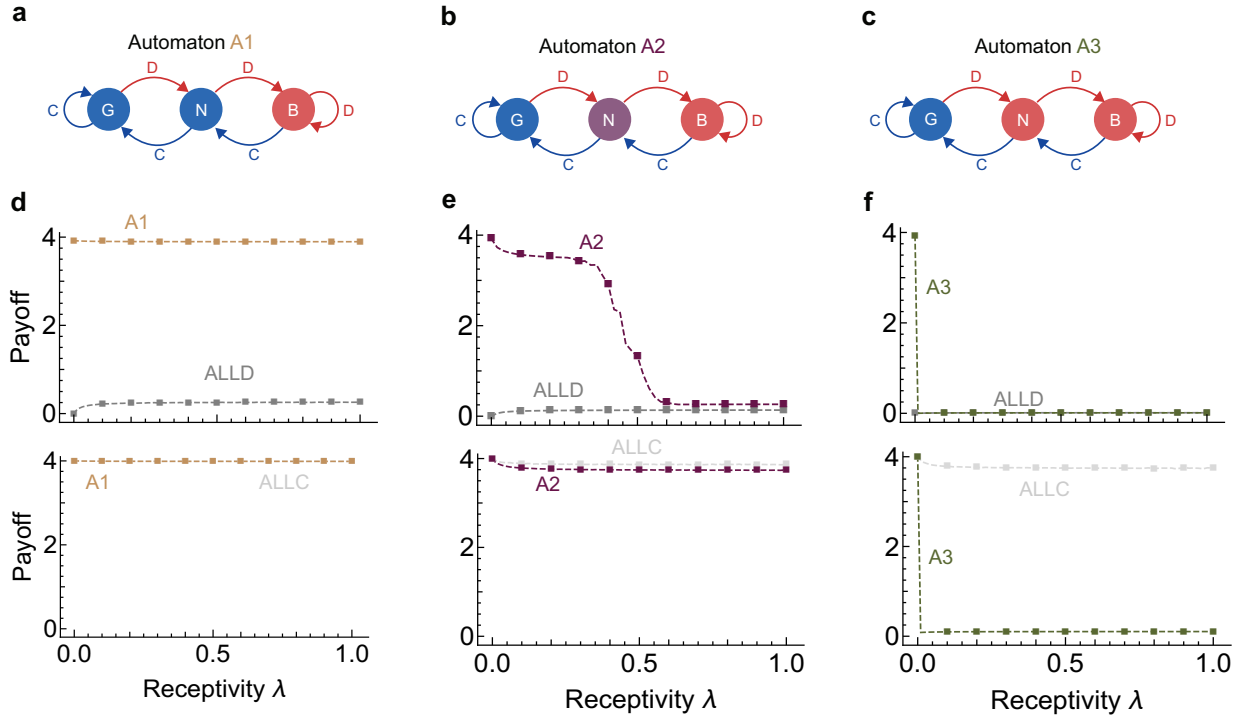
Extended Data Figure 3: Impact of different model parameters on the co-evolution of direct and indirect reciprocity. We show how our evolutionary results in **Fig. 5** are affected as we change different parameters of our model. In each panel, we vary one parameter and leave all others constant. We consider the same three scenarios as in **Fig. 5a–c**: few interactions and unreliable information (blue), intermediate interactions and reliable information (orange), and many interactions and unreliable information (green). We employ two complementary simulation techniques. In the upper half, each data point represents the average of a single simulation. This simulation was run for sufficiently long such that the averages converge and are independent of the initial condition. This typically happens after 10^7 mutant strategies have been introduced into the population. In the lower panels, each data point represents the average of 200 simulations with a random initial population. Here, each simulation only introduces 10^5 mutant strategies. For the parameters, we consider variation in the benefit-to-cost ratio (**a,b**), the population size (**c,d**), the selection strength (**e,f**), and the mutation rate (**g,h**). Our simulations suggest that each of these parameters can have a considerable impact on the evolving cooperation rates and the player’s propensity to adopt indirect reciprocity. For example, for the orange curve in panel **e**, we observe that the effect of selection strength on cooperation can be non-monotonic. We further discuss these dependencies in **Extended Data Fig. 4** and **SI Section 5**. In general, however, we recover the following regularities from **Fig. 5**: (i) Substantial cooperation only evolves in the second and third scenario (i.e., for the cooperation rates, the blue curve is systematically below the other curves). (ii) If cooperation evolves, players prefer indirect reciprocity when there are intermediately many interactions and outside information is reliable. They prefer direct reciprocity when there are many interactions and when outside information is noisy (i.e., for the proportion of indirect reciprocity, the orange curve is systematically above the green curve).



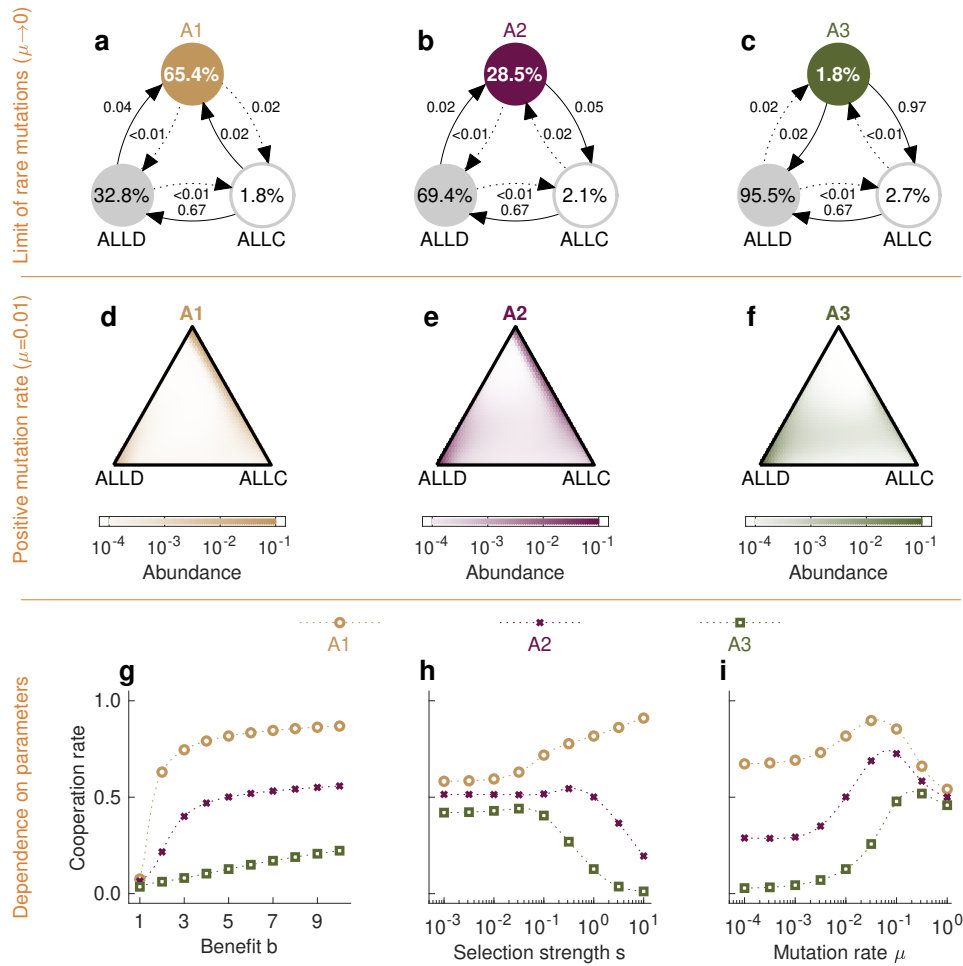
Extended Data Figure 4: Impact of selection strength on indirect reciprocity. As shown in the upper panel of **Extended Data Fig. 3e**, selection can sometimes have a non-monotonic effect on cooperation. For intermediate interactions and reliable information ($\delta = 0.9, \varepsilon = 0.001$, depicted by the orange curve in **Extended Data Fig. 3e**), we have observed that the evolving cooperation rate is 53.4% for $\beta = 1$, increases to 77.3% for $\beta = 10$, and reduces to 61.5% for $\beta = 100$. Here we present additional simulations to shed further light on this non-monotonicity. **a,b**, We considered initial resident populations that either adopt a defective strategy or a conditionally cooperative strategy. We recorded how long it takes the evolutionary process until the resident strategy is replaced, and what the cooperation rate of the invading strategy is. Dots show the outcome of individual simulations, and the curves represent averages. The results suggest that the non-monotonicity of cooperation is not due to a reduced stability of cooperative strategies. They remain highly robust even for large selection strengths. Moreover, when selection is strong, they are typically invaded only by other cooperative strategies. **c,d** In a next step, we recorded the distribution of cooperation over time for three different selection strengths for the process considered in **Extended Data Fig. 3e**. We find that this distribution becomes more extreme with increasing selection strength: individuals either become highly cooperative or highly non-cooperative. However, the proportion of non-cooperative populations grows faster than the proportion of cooperative populations.



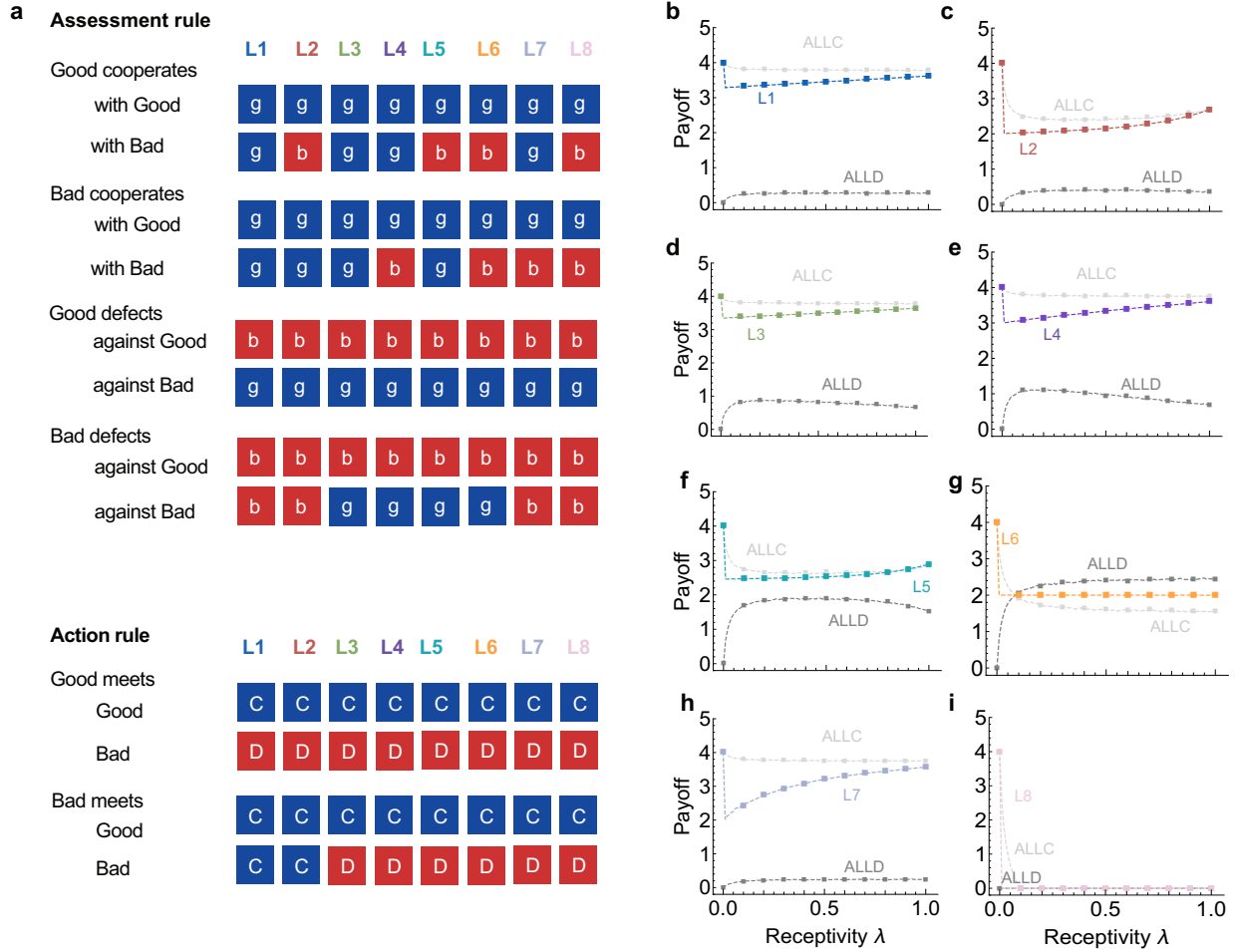
Extended Data Figure 5: Effect of different types of errors and incomplete information on cooperation. **a**, To explore how sensitive our results are to different kinds of errors and incomplete information, we have repeated the rare mutation simulations shown in **Fig. 5d**, reproduced here. **b**, While the baseline model assumes that only indirect observations are subject to perception errors, here we explore the effects when direct observations are equally prone to errors. We find that cooperation is substantially reduced compared to the baseline scenario. Moreover, direct reciprocity is only favoured for very large continuation probabilities. **c**, We have also explored the effect of additional implementation errors on cooperation. To this end, we assume here that players mis-implement their intended action with fixed probability $e = 0.01$. Compared to the baseline model without such errors, we find that there is less cooperation and less direct reciprocity. **d**, To mimic the dynamics that arises when defectors strategically conceal their bad actions, we have also considered a model in which defective actions are misperceived with probability ε , whereas cooperative actions are always observed faithfully. Because this assumption reduces the total rate at which errors occur compared to the baseline scenario, we observe more cooperation and players are more reliant on indirect reciprocity. **e**, Here we assume that individuals observe third-party interactions only with probability $\nu = 0.01$. Due to the scarcity of information, players who take any third-party information into account are almost indistinguishable from those players who do not. As a result, cooperation is largely independent of observation errors, and the region in which indirect reciprocity is favoured has vanished. Unless noted otherwise, all parameters are the same as in **Fig. 5d**.



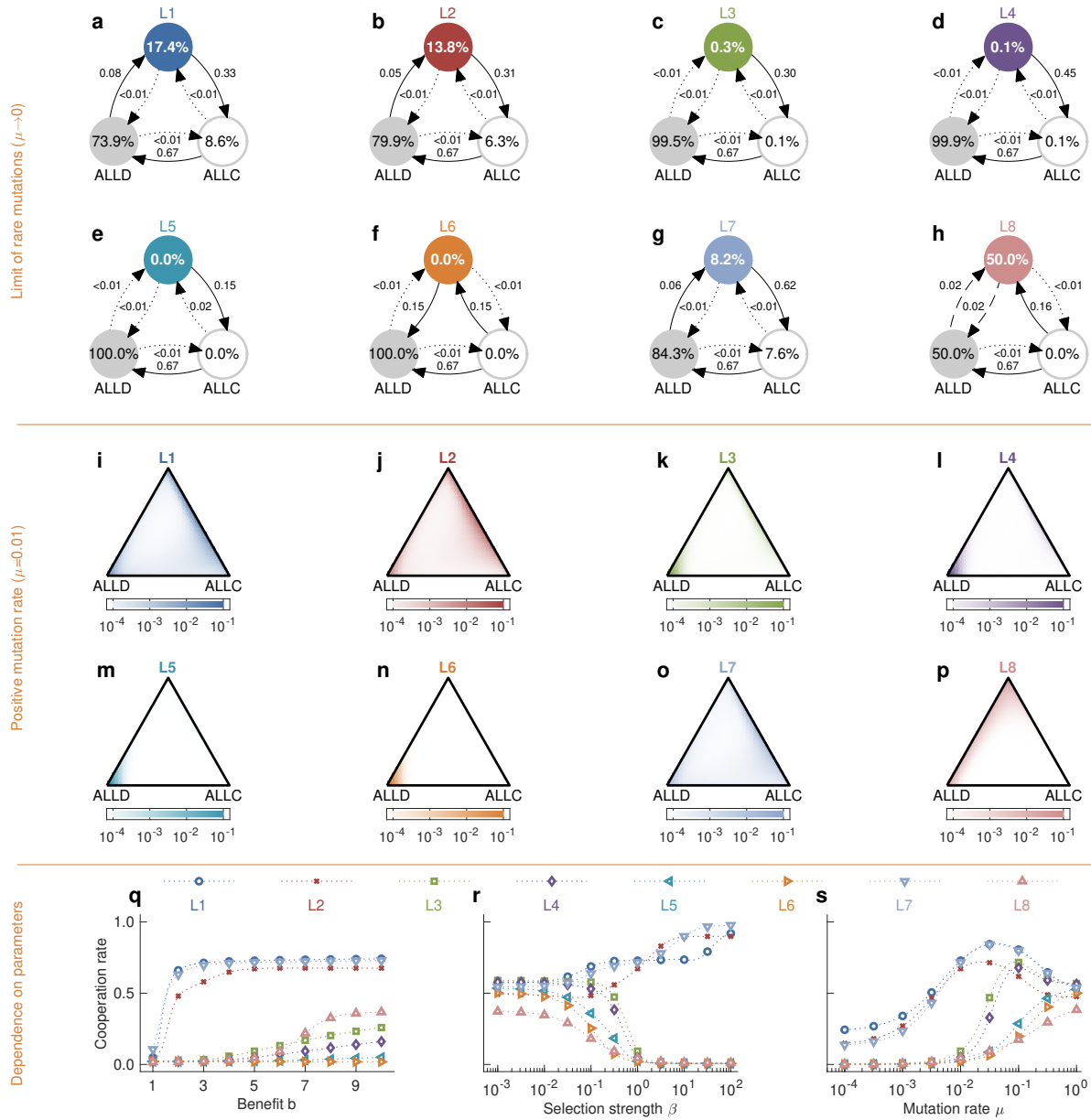
Extended Data Figure 6: Direct and indirect reciprocity for finite-state automata with three states. In an extension of our model, we allow players to assign more nuanced reputations to their co-players. We illustrate this approach by considering finite state automata with three states - good (G), neutral (N) and bad (B), with G as the initial state. We assume $n-1$ residents employ the respective finite-state automaton strategy, while the remaining player uses either *ALLC* or *ALLD*. We simulate the players' payoffs for various values of $\lambda \in [0, 1]$. We consider three different automaton strategies employed by the residents. The automata differ in how they deal with co-players that are assigned a neutral reputation. **a**, Players with the first automaton $A1$ are fully cooperative when they encounter a co-player with neutral reputation. This strategy can sustain cooperation among itself. However, a single *ALLC* player obtains approximately the same payoff as the residents, and hence can invade by (almost) neutral drift. **b**, According to the second automaton $A2$, players cooperate against neutral opponents with 50% probability. This strategy can be invaded by *ALLC* for all $\lambda > 0$. **c**, According to $A3$, players defect against co-players with a neutral reputation. This strategy is not stable against *ALLC* for $\lambda > 0$, and residents fail to cooperate with each other altogether.



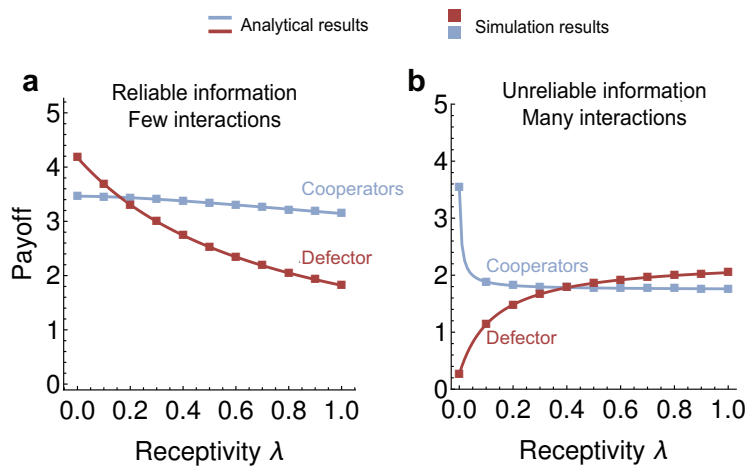
Extended Data Figure 7: Evolutionary competition between finite state automata, ALLC, and ALLD. We have explored the evolutionary dynamics when population members can choose between *ALLC*, *ALLD*, and one of the three finite-state automata introduced in **Extended Data Fig. 6**, and a fixed receptivity $\lambda = 0.1$. The numbers in each circle denote how often the respective strategy is played on average. Arrows illustrate how likely a single mutant fixes in the respective resident population. Solid arrows indicate that the fixation probability is larger than the neutral $1/n$, whereas for dotted arrows this probability is smaller than neutral. We find that only the first automaton A_1 can outperform both *ALLC* and *ALLD*. **d–f**, In a next step, we have explored the same scenario for a positive mutation rate $\mu = 0.01$. The triangles represent the possible population compositions. Each corner corresponds to a homogeneous population, whereas the center corresponds to a perfectly mixed population. The color code reflects how often we observe the respective population composition over the course of evolution. We find that most of the time, populations are either in the neighborhood of *ALLD*, or they represent some mixture between the automaton strategy and *ALLC*. **g–i**, We have re-run the simulations in panels **d–f**, but now varying either the benefit of cooperation, the selection strength, or the mutation rate. In all cases, we observe that the first automaton is most favorable to cooperation. Interestingly, we observe the largest cooperation rate for intermediate mutation rates. This result, however, may be due to the fact that players can only choose from an unbalanced strategy space, as discussed in detail in **SI Section 6.3**.



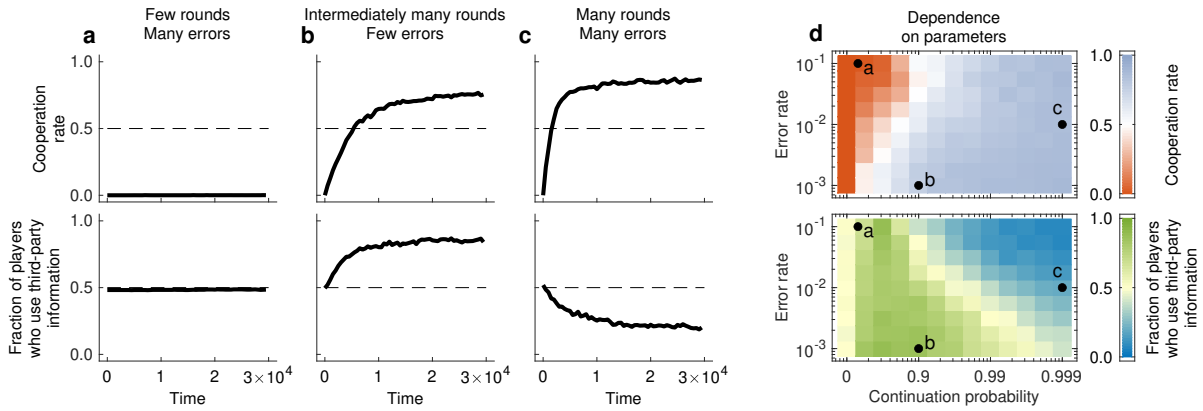
Extended Data Figure 8: Performance of leading-eight strategies under direct and indirect reciprocity. **a**, Previous research has suggested that there are eight stable third-order strategies of indirect reciprocity that can sustain cooperation²², called the leading eight, $L1-L8$. They consist of two components, an assessment rule and an action rule. The assessment rule determines how players evaluate each other's actions, depending on the previous reputations of the involved players. The action rule determines how to interact in the game, depending on one's own reputation and on the reputation of the co-player. **b-i**, To explore the stability of these strategies, we consider a population in which $n-1$ players adopt one of the leading-eight strategies. The remaining player either adopts *ALLC* or *ALLD*. Our results for $\lambda > 0$ reflect previous findings³²: in the presence of perception errors, all leading-eight strategies are susceptible to invasion by either *ALLC* or *ALLD*. Only for $\lambda = 0$ (when perception errors are absent), the leading-eight strategies are stable against both mutant strategies.



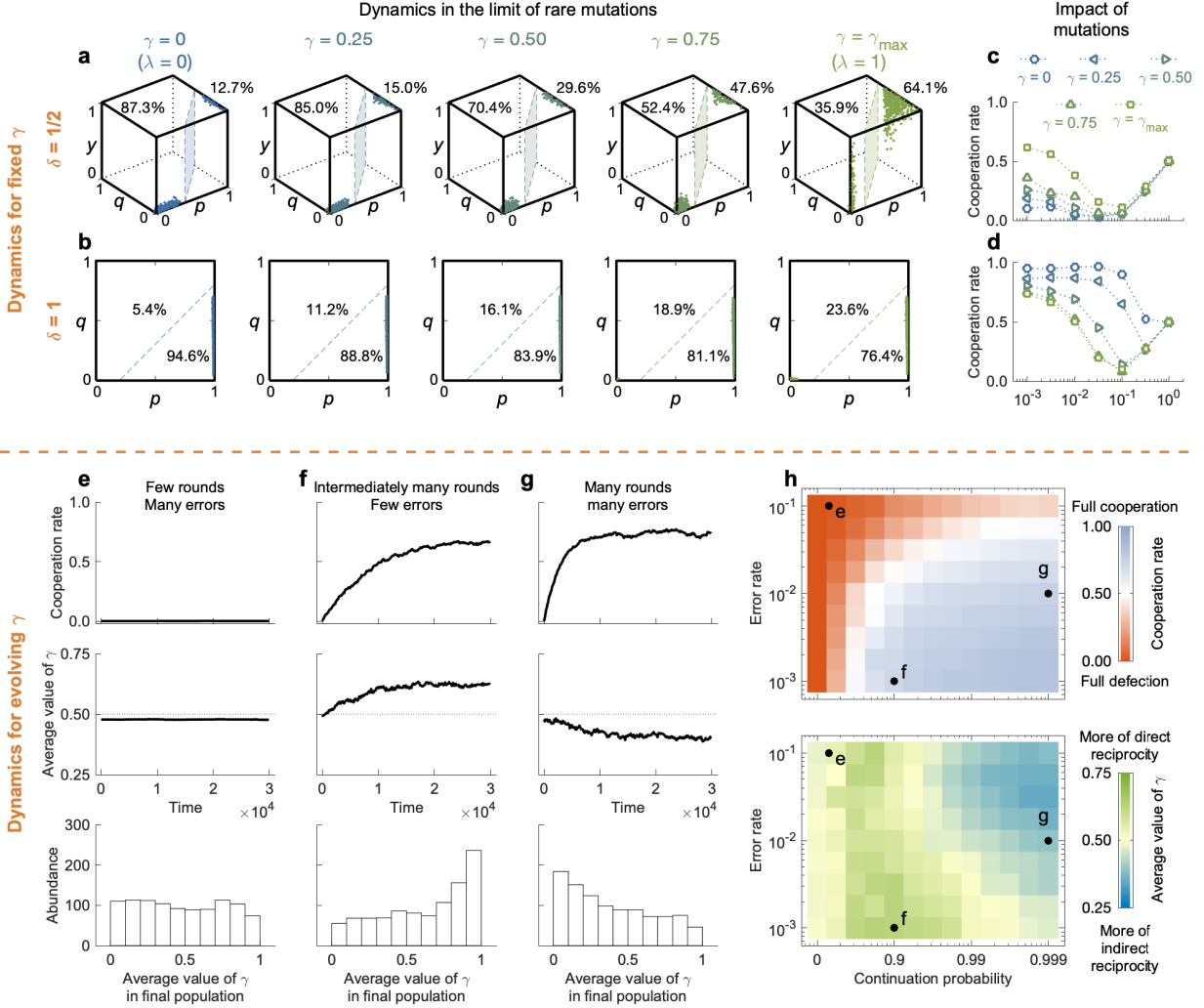
Extended Data Figure 9: Evolutionary dynamics of the leading-eight. Similar to **Extended Data Fig. 7** for finite state automata, this figure explores how each of the leading-eight fares in an evolutionary competition against *ALLC* and *ALLD* for a fixed receptivity $\lambda = 0.1$. **a–h**, When mutations are rare, only ‘Judging’ (L_8) is played in notable proportions. However, in the presence of perception errors, this strategy tends to assign a bad reputation to other players with the same strategy, such that everyone defects eventually³². **i–p**, When mutations are more common, some of the leading-eight strategies can stably coexist with *ALLC*. We observe such cooperative coexistences for L_1 , L_2 , and L_7 . **q–s**, These three strategies also yield substantial cooperation rates when we vary the benefit of cooperation, the selection strength, and the mutation rate. With respect to mutation, we again observe that intermediate mutation rates are most favorable to cooperation. However, this finding may not be robust, because the strategy space is again unbalanced. For a more detailed discussion, see **SI Section 6.4**.



Extended Data Figure 10: Simulations of the game dynamics confirm the results of the analytical payoff calculations. Here, we explore whether equation (7) gives an accurate prediction of the resulting payoffs when all players adopt some (reactive) strategy (y, p, q, λ) . To this end, we consider $n - 1$ conditional cooperators with strategy $(1, 1, 0.01, \lambda)$. The remaining player is a defector. We calculate payoffs in two different ways, by using the formula (7), and by simulating the game dynamics explicitly. **a**, When errors are rare and the continuation probability is comparably small, cooperators can only outperform defectors when they take indirect information into account. **b**, In contrast, when information is noisy and there are many pairwise interactions, cooperators obtain a better payoff when they ignore indirect information. In both cases, our analytical results agree with the simulations.



Extended Data Figure 11: The coevolution of conditional cooperation and information use in a model of “purified” indirect reciprocity. For the model discussed in the main text, we assume that after a direct interaction, players always update their co-player’s reputation. After an indirect interaction, they update their co-player’s reputation with probability λ . These assumptions imply that even if $\lambda = 1$, players may occasionally react based on their direct experiences. This happens, for example, if two players have two consecutive interactions. In that case, the players’ behavior in the second round will depend on the outcome of the first. While such an assumption seems realistic, it can be useful to compare purified versions of both kinds of reciprocity. We explore such a model in **SI Section 6.2**. There, the players’ strategies take the form (y, p, q, κ) . A value of $\kappa = 0$ means players ignore all of a co-player’s third party interactions (similar to the case $\lambda = 0$ discussed before). A value of $\kappa = 1$, however, now means that players ignore all direct information they may have. In the limit of large populations, the two strategy spaces (y, p, q, λ) and (y, p, q, κ) yield identical results. Here we repeat the simulations performed in **Fig. 5a–d** for a finite population of $n = 50$, using the alternative strategy space (y, p, q, κ) . The results are almost indistinguishable from the results shown in **Fig. 5a–d**.



Extended Data Figure 12: Evolution of cooperation for players with intermediate degrees of receptivity. In the main text figures **Fig. 3–Fig. 5**, we explore situations in which individuals can choose strategies where they either only take direct information into account ($\lambda = 0$), or where they take all information into account ($\lambda = 1$). Here we repeat these simulations in a setup where intermediate values of λ are permitted. To this end, we define a quantity γ . This quantity is the probability that a player’s decision is based on the co-player’s behavior towards third parties, see Eq. (10) in **Methods**. For $0 \leq \lambda \leq 1$ we obtain $0 \leq \gamma \leq \gamma_{\max} := (n-2)/(n-1)$. **a,b**, We repeat the simulations in **Fig. 3a,b** for various values of γ . We observe that cooperation is never most likely to evolve for intermediate values of γ . Either most cooperation evolves for $\gamma = \gamma_{\max}$ (in panel **a**), or for $\gamma = 0$ (in panel **b**). **c,d**, Similarly, we repeat the simulations in **Fig. 4d,f** for various values of γ . Again, the average cooperation rates for intermediate γ are strictly in between the results for $\gamma = 0$ and $\gamma = \gamma_{\max}$. **e–h**, Finally, we repeat the simulations shown in **Fig. 5a–d**, allowing for mutant strategies (y, p, q, λ) that lead to arbitrary values of γ between 0 and γ_{\max} . Especially for larger error rates, we observe that the evolving cooperation rates are now smaller. Nevertheless, the general patterns of **Fig. 5** remain: (i) When there are only few rounds and many observation errors, cooperation does not evolve. (ii) When there are intermediately many rounds and few errors, cooperation evolves and players tend to put more weight on indirect information (that is, γ tends to be larger than 1/2). In particular, strategies with $\gamma \approx \gamma_{\max}$ are most abundant. (iii) When there are many rounds and intermediately many errors, cooperation evolves and players tend to put more weight on direct information. Here, players are most likely to adopt a strategy with $\gamma \approx 0$. See **SI Section 5.4** for details.