



# Computational evolution of social norms in well-mixed and group-structured populations

Yohsuke Murase<sup>a,1</sup> and Christian Hilbe<sup>b</sup>

Edited by Marcus Feldman, Stanford University, Stanford, CA; received April 5, 2024; accepted July 15, 2024

Models of indirect reciprocity study how social norms promote cooperation. In these models, cooperative individuals build up a positive reputation, which in turn helps them in their future interactions. The exact reputational benefits of cooperation depend on the norm in place, which may change over time. Previous research focused on the stability of social norms. Much less is known about how social norms initially evolve when competing with many others. A comprehensive evolutionary analysis, however, has been difficult. Even among the comparably simple space of so-called third-order norms, there are thousands of possibilities, each one inducing its own reputation dynamics. To address this challenge, we use large-scale computer simulations. We study the reputation dynamics of each third-order norm and all evolutionary transitions between them. In contrast to established work with only a handful of norms, we find that cooperation is hard to maintain in well-mixed populations. However, within group-structured populations, cooperation can emerge. The most successful norm in our simulations is particularly simple. It regards cooperation as universally positive, and defection as usually negative—unless defection takes the form of justified punishment. This research sheds light on the complex interplay of social norms, their induced reputation dynamics, and population structure.

social norms | indirect reciprocity | evolution of cooperation | evolutionary game theory

Humans exhibit a remarkable capacity for cooperation (1). We cooperate with friends and family members, but also with strangers we are unlikely to ever meet again (2). When people cooperate beyond their own social circle, they often follow social norms. These norms can be seen as a set of simple principles that tell people how they should act, and how they should assess other people's actions (3, 4). Importantly, such norms can be subject to evolutionary change. They persist as long as individuals have an incentive to respect them. They get replaced once individuals find profitable ways to deviate. The question of which social norms are stable and which lead to cooperation has given rise to an active research area called indirect reciprocity (5–12).

To get a first quantitative understanding of social norms, it has become common to study models with two possible reputations only (13–15). Individuals can either be, say, “good” or “bad.” One purpose of a social norm, then, is to help people classify who they should regard as good. Depending on the social norm's complexity, there are different ways to do so (16). In the simplest case of a “first-order norm”, reputations only depend on what an individual did. For example, according to the norm “image-scoring” (17), cooperators should be regarded as good and defectors as bad. Although this norm is a reasonable first approximation of human morality (18), it is unstable (19, 20). Stable norms do not only need to incentivize individuals to cooperate. They also need to allow them to defect against bad interaction partners. This logic is incorporated in “second-order norms”. Here, reputations do not only depend on what an individual did, but also to whom. One example is “Stern Judging” (21). Here, individuals are regarded as good as long as they cooperate with well-reputed group members and defect against the bad ones. Next comes the class of “third-order norms”. Here, people also take into account the focal individual's previous reputation when assigning a new one. This hierarchy of social norms can be further extended ad infinitum (22).

Whether or not a social norm is stable, however, does not only depend on its complexity. It also depends on how information spreads in a population, and how individuals form their opinions. Traditionally, many models study the case of “public assessment” (21–38). These models assume that people unanimously agree on which reputation should be assigned to each other. For example, if Alice thinks Bob deserves a good reputation, then so does Carol, Dave, and Eve (and anyone else). The assumption of public information is useful because it makes models more mathematically tractable. This allowed Ohtsuki & Iwasa to fully characterize a class of third-order norms that

## Significance

Social norms have a profound impact on our daily lives. They determine how we behave and how our behaviors are evaluated by others. The literature on indirect reciprocity explores how social norms for cooperation emerge and propagate. A systematic evolutionary analysis of all norms, however, has been difficult, due to the enormous number of possible norms and pathways between them. Here, we use a supercomputer to study transitions between thousands of norms. We find that once all norms of a given complexity class are permitted, cooperation in well-mixed populations is surprisingly fragile. Yet when populations have some natural group structure, cooperative norms can persist. Our findings show how social norms, reputations, and population structure interact to promote cooperation.

Author affiliations: <sup>a</sup>RIKEN Center for Computational Science, Kobe 650-0047, Japan; and <sup>b</sup>Max Planck Research Group Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, Plön 24306, Germany

Author contributions: Y.M. and C.H. designed research; Y.M. performed research; Y.M. and C.H. analyzed data; and Y.M. and C.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

<sup>1</sup>To whom correspondence may be addressed. Email: yohsuke.murase@gmail.com.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2406885121/-/DCSupplemental>.

Published August 8, 2024.

can sustain full cooperation, called the “leading-eight” (23). However, the assumption of public assessment can be overly restrictive. After all, individuals may have access to different pieces of information, they may misinterpret a given interaction, or they may apply different social norms altogether. As a result, they may hold different views. In corresponding models with “private assessment” (39–52), Alice might think of Bob as good even though both Carol and Dave disagree. Mathematically, these models are more challenging. While the individuals’ opinions might still be correlated, they are no longer correlated perfectly. As a consequence, analytical results only exist in certain limiting cases (51–57).

To date, there are two main approaches to study models with private assessment (for a detailed overview of the literature, see *SI Appendix*). The first approach is a local stability analysis. Here, researchers consider a resident population with a given norm, and test its stability against an (infinitesimally) small number of deviating “mutants” (42, 57). This approach allows for analytical results, but it is essentially static. The other approach allows mutants to become more common, and to eventually replace the resident norm (45–54). This approach, however, typically considers a restricted set of possible mutant norms. In many cases, the set of mutant norms only contains the most extreme possibilities, “Always Cooperate (ALLC)” and “Always Defect (ALLD)”. A more complete view of the evolution of social norms would require a global analysis that includes all possible norms of a given complexity class. So far, however, a global analysis has been infeasible. Even for the comparably simple space of third-order norms, there are  $K = 2,080$  distinct possibilities. Taking into account all possible transitions between them, in a population of size  $N$ , gives rise to approximately  $NK^2$  different cases. Each case requires its own simulation to compute the resulting reputation dynamics. So far, the respective computational costs have been prohibitive.

Herein, we report the results of such a global analysis, obtained by running the necessary simulations on a supercomputer (*Materials and Methods*). In contrast to previous studies with a limited set of mutant norms, we find that in well-mixed populations, cooperation is unlikely to emerge. This result, however, does not mean that cooperation through indirect reciprocity is impossible altogether. As a proof of principle, we also study social interactions in group-structured populations. For those interactions, we find that cooperation evolves reliably. Our analysis highlights the importance of one particular leading-eight norm, L1, which previously received little attention. This norm is similar to image scoring, but with one crucial difference: It permits good individuals to defect against bad group members, without harming their own reputation. Overall, these results highlight how the interplay of social norms and population structure can facilitate the emergence of indirect reciprocity.

## Model

**Model of Indirect Reciprocity.** Our analysis follows an established framework for indirect reciprocity (46–48). We consider a well-mixed population of size  $N$ , as illustrated in Fig. 1A. The population members are referred to as players. Players engage in a sequence of donation games as follows. Each round, two players are randomly drawn from the population, one as a donor and the other as a recipient. The donor decides whether or not to provide a benefit  $b$  to the recipient at a personal cost of  $c$ , with  $0 < c < b$ . These two possible actions are referred to as cooperation ( $C$ , providing the benefit) and defection

( $D$ , doing nothing). This process is iterated over many rounds, by consecutively drawing random donor-recipient pairs.

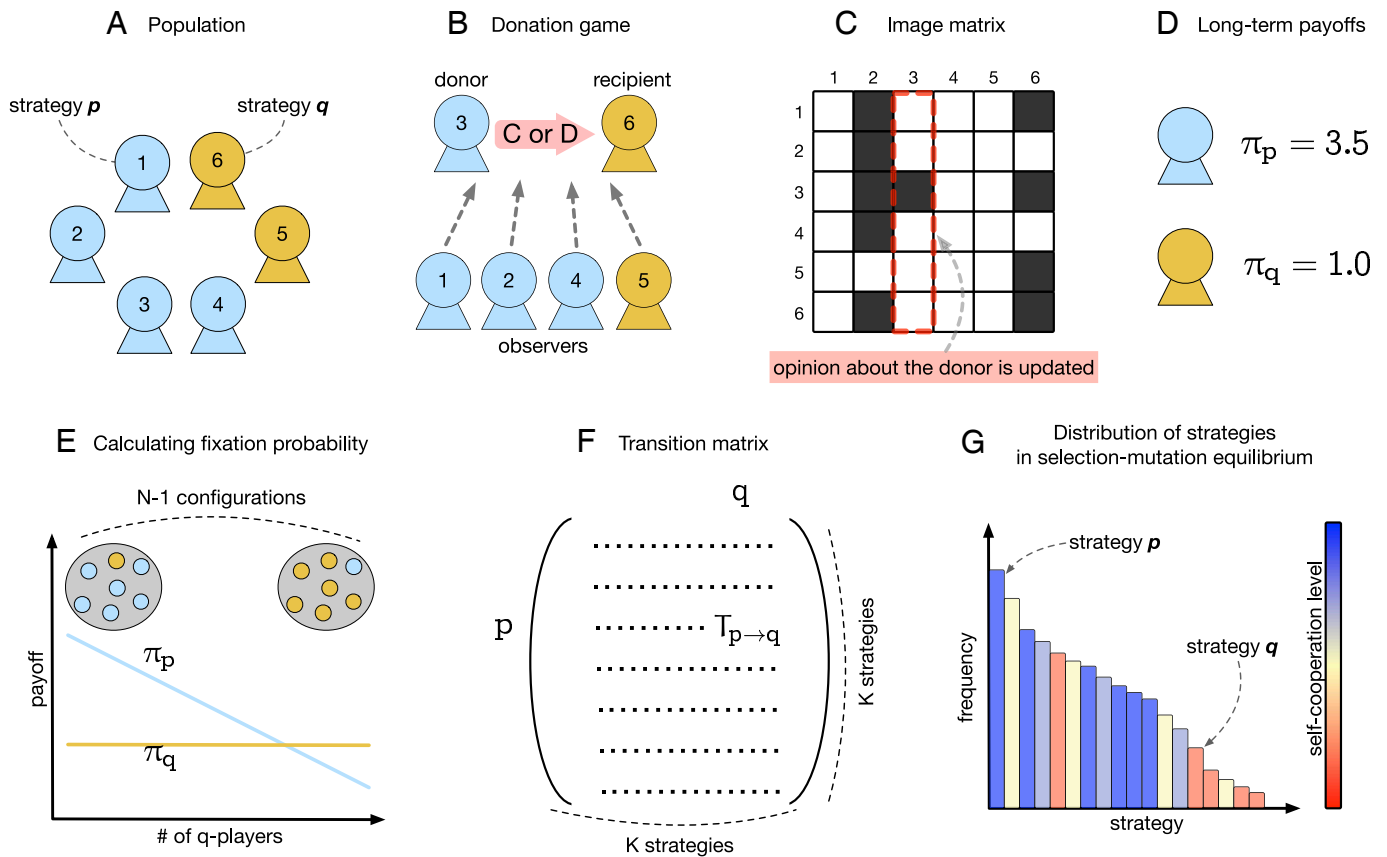
During this sequence of donation games, players form opinions about each other. The opinion player  $i$  holds about  $j$  is denoted as  $m_{ij}$  (players have an opinion about everyone, including themselves). Opinions are good ( $G$ ) or bad ( $B$ ), and they can change over time. Each donation game is observed by all other population members simultaneously. Accordingly, we speak of a “simultaneous observation model” (results for a more general model are in *SI Appendix*). Based on their respective information, each observer independently assesses the donor. The way observers make these assessments depends on their assessment rule ( $R$ ). The way donors decide whether to cooperate in the donation game is determined by their action rule ( $P$ ). Both of these rules are specified in more detail below. Importantly, different players may use different action and assessment rules. In the following, we refer to a combination of an action and assessment rule as a player’s strategy (the term “social norm” is often used synonymously; in the following, we prefer to use “strategy” to indicate that this is a property of the individual, not of the population).

**Action and Assessment Rules.** Herein, individuals make their decisions based on third-order strategies (22, 23). The first component of a strategy, its action rule, determines whether or not donors cooperate with a given recipient. For third-order strategies, this decision may depend on both the donor’s and the recipient’s reputation. As a result, when donor  $i$  is matched with recipient  $j$ , the donor’s action  $A_{ij} \in \{C, D\}$  depends on both  $m_{ii}$  and  $m_{ij}$ ,  $A_{ij} = P_i(m_{ii}, m_{ij})$ . As an example, suppose the donor wishes to cooperate if and only if she deems the recipient as good. In that case,  $P_i(m_{ii}, G) = C$  and  $P_i(m_{ii}, B) = D$ , irrespective of the donor’s self-image  $m_{ii}$ .

The donor’s decision is observed by all other population members (Fig. 1B). Each observer  $k$  then independently updates their opinion about donor  $i$  according to the other component of the strategy, the assessment rule. The updated opinion  $m'_{ki} \in \{G, B\}$  depends on the observer’s previous opinion  $m_{ki}$  of the donor, her previous opinion  $m_{kj}$  of the recipient, and the donor’s action  $A_{ij}$ . That is,  $m'_{ki} = R_k(m_{ki}, m_{kj}, A_{ij})$ . For example, according to image scoring, a player’s assessment only depends on the donor’s action. Cooperation is deemed as good,  $R_k(m_{ki}, m_{kj}, C) = G$ , and defection is evaluated as bad,  $R_k(m_{ki}, m_{kj}, D) = B$  for all  $m_{ki}, m_{kj}$ . In general, when two observers differ in their initial assessment of a donor, they may also disagree on the donor’s updated assessment, even if both use the same strategy and observe the same interaction. For a few examples of important strategies, Table 1 defines the “leading-eight.” This is a set of strategies that can maintain cooperation in a public information setting (23).

Since  $G$  and  $B$  are merely two labels, strategies remain unchanged if we invert the meaning of these labels (23). Taking this symmetry into account, there are 2,080 distinct third-order strategies (*Materials and Methods*). We denote that set by  $\mathcal{S}_3$ . For comparison, we also consider the set of the second-order rules,  $\mathcal{S}_2$ , whose size is 36. For those norms, the action rule only depends on the donor’s opinion about the recipient. In addition, the assessment rule only depends on the observer’s opinion about the recipient, and on the donor’s action. Among the leading-eight, L3 (Simple Standing) and L6 (Stern Judging) are the only second-order strategies.

Herein, we define ALLD as the class of all strategies that always defect regardless of reputation,  $P(X, Y) = D$  for any  $(X, Y)$ .



**Fig. 1.** A schematic overview of our computational framework to study indirect reciprocity. (A) We consider a population of players. Different players may adopt different strategies (social norms). That is, they may differ in how they evaluate each other's actions, and how they decide with whom to cooperate. In the short run, these strategies are fixed. (B) Players engage in a series of donation games. Each round, a donor and a recipient are randomly drawn. The donor decides whether to cooperate or to defect. All population members observe the donor's decision. Based on these observations, they update their opinion of the donor. (C) The population's current state can be summarized by an image matrix. For each pair of players  $i$  (row) and  $j$  (column), this matrix describes  $i$ 's opinion about  $j$ . This opinion can be either good (white) or bad (black). (D) Based on the dynamics of this image matrix, we compute the expected payoff of each strategy. (E) In the long run, players' strategies are allowed to change. We assume the change of strategies follows an evolutionary process based on imitation events and mutations. (F) Here, we present results when mutations are rare. In that case, the evolutionary dynamics can be described by a Markov chain. The transitions of this Markov chain correspond to the transitions from one homogeneous population to another. (G) By computing the invariant distribution of this Markov chain, we assess how often each strategy is played in the long run.

All these strategies behave identically, independent of their respective assessment rules. There are 136 and 10 ALLD strategies in  $\mathcal{S}_3$  and  $\mathcal{S}_2$ , respectively. Similarly, there is the same number of ALLC strategies, which always cooperate regardless of reputations,  $P(X, Y) = C$  for any  $(X, Y)$ .

**Reputation Dynamics.** In the short run, we take each player's strategy to be fixed. Given the players' strategies and their initial reputations, we can describe the current state of the population by an image matrix  $M = (m_{ij})$ , see Fig. 1C. At any given round, this matrix summarizes the opinions that players have of each other

**Table 1. The leading-eight strategies describe a set of third-order strategies that can maintain cooperation under public assessment (23)**

	(G, G)			(G, B)			(B, G)			(B, B)		
	P	R(C)	R(D)	P	R(C)	R(D)	P	R(C)	R(D)	P	R(C)	R(D)
L1	C	G	B	D	<b>G</b>	G	C	G	B	<b>C</b>	<b>G</b>	<b>B</b>
L2 (Consistent Standing)	C	G	B	D	<b>B</b>	G	C	G	B	<b>C</b>	<b>G</b>	<b>B</b>
L3 (Simple Standing)	C	G	B	D	<b>G</b>	G	C	G	B	<b>D</b>	<b>G</b>	<b>G</b>
L4	C	G	B	D	<b>G</b>	G	C	G	B	<b>D</b>	<b>B</b>	<b>G</b>
L5	C	G	B	D	<b>B</b>	G	C	G	B	<b>D</b>	<b>G</b>	<b>G</b>
L6 (Stern Judging)	C	G	B	D	<b>B</b>	G	C	G	B	<b>D</b>	<b>B</b>	<b>G</b>
L7 (Staying)	C	G	B	D	<b>G</b>	G	C	G	B	<b>D</b>	<b>B</b>	<b>B</b>
L8 (Judging)	C	G	B	D	<b>B</b>	G	C	G	B	<b>D</b>	<b>B</b>	<b>B</b>

In this table, the *Top* row  $(X, Y)$  indicates the reputations of the donor and the recipient, respectively. For instance,  $(G, B)$  means a good donor meets a bad recipient. The rules  $P, R_1(C), R_1(D)$  indicate the prescribed action, the assessment when cooperation (C) is observed, and the assessment when defection (D) is observed, respectively. The entries that are different from each other are highlighted in bold.

(i.e., the reputations they assign to each other). The image matrix changes after each donation game. The new matrix depends on which player has been chosen as the donor, the donor's action, and how this action is assessed by all other population members. Here, we allow assessments to be subject to errors: An observer who would usually assign a good reputation to a given donor assigns a bad one with probability  $\epsilon$  (and vice versa). For  $\epsilon > 0$ , the reputation dynamics becomes ergodic. As a result, the reputation dynamics reaches a stationary state after sufficiently many rounds. This long-run dynamics is independent of the players' initial reputations. This allows us to compute how often people deem each other as good in the long run, and how often they cooperate with each other. Based on this, we compute the players' average payoffs over the entire sequence of donation games, see Fig. 1D.

**Evolutionary Dynamics.** On a larger time scale, players can change their strategies. To model the evolution of strategies, we use a pairwise comparison process (58). In each time step of the process, one player is randomly chosen. With probability  $\mu$ , a mutation occurs. In that case, the player adopts a random strategy, uniformly among all strategies in  $\mathcal{S}_3$  (or  $\mathcal{S}_2$ , respectively). With the remaining probability  $1-\mu$  the focal player  $i$  chooses a random role model  $j$  from the population. Player  $i$  then imitates  $j$ 's strategy with probability  $\varphi_{i \rightarrow j}^{\text{in}}$ , which is an increasing function of the payoff difference between  $j$  and  $i$  (*Materials and Methods*). In particular, the more profitable  $j$ 's strategy, the more likely  $j$  is imitated.

For our evolutionary simulations, we consider the low mutation rate limit (59, 60). That is, mutations are rare enough such that typically only a single resident strategy is present in the population. Only occasionally, a mutant appears. This mutant reaches fixation or goes extinct before the next mutant is introduced. The mutant's fixation probability can be computed explicitly (61). It depends on the players' payoffs for any possible population configuration consisting of the mutant and the resident strategy (Fig. 1E). The overall evolutionary dynamics among all strategies can then be described by a Markov process. The possible states correspond to the different homogeneous resident populations. The transition probabilities from one state to another depend on the respective fixation probabilities (Fig. 1F). By computing the stationary distribution of this Markov process, we can calculate how frequently each strategy is played in the long run (Fig. 1G).

When only three different strategies compete (as in many previous studies), it is straightforward to calculate the stationary distribution (e.g. ref. 46). However, once all strategies in  $\mathcal{S}_3$  compete, the situation is more challenging. Now, we need to calculate fixation probabilities for all combinations of third-order strategies. Given there are  $K = 2,080$  such strategies, the number of pairwise fixation probabilities is  $K(K-1)$ . For each of these, we need to calculate the long-term payoffs of the respective residents and mutants for  $N-1$  different population configurations (as illustrated in Fig. 1E). For each configuration, we need to simulate the reputation dynamics for a sufficiently long time. Here, we use  $N = 50$  and take  $n = 5$  independent samples for each parameter. Overall, we conduct  $nK(K-1)(N-1)/2 \approx 5.5 \times 10^8$  simulations, each one recording the outcome of  $10^6$  donation games. To deal with these computational requirements, all calculations are parallelized and run on a supercomputer (*Materials and Methods*). In the following sections, we present our findings for a specific set of parameters. Results for other parameter values are qualitatively similar, as shown in *SI Appendix*.

## Results

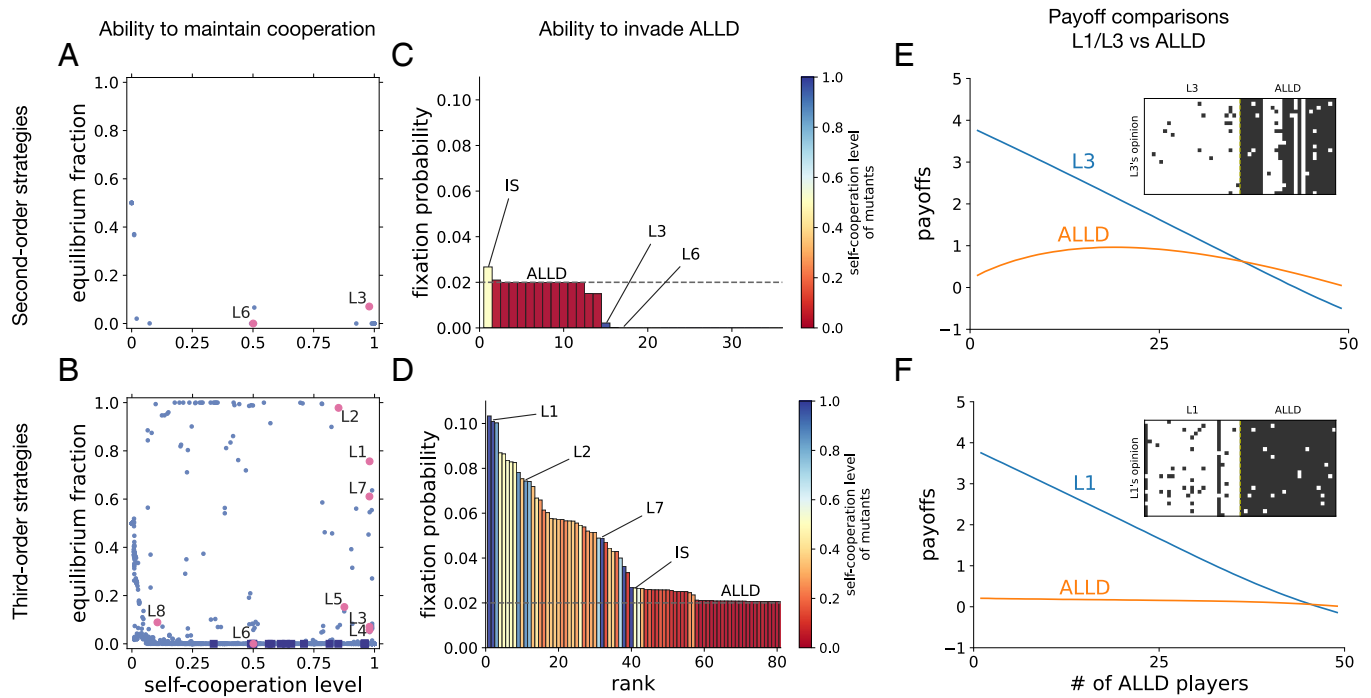
**Evolutionary Dynamics in Three-Strategy Systems.** To get a first overview of the evolutionary performance of different strategies, we study a restricted setup. Following earlier work (45–54), we explore different three-strategy systems, where a focal strategy  $X$  competes with ALLC and ALLD. In each case, we compute  $X$ 's self-cooperation level  $p_c(X)$  (the cooperation level of a homogeneous  $X$  population). In addition, we compute the strategy's abundance  $f_{\text{eq}}^\Delta(X)$  in the resulting stationary distribution. If  $X$  is to maintain cooperation, both quantities should be close to one. However, our simulations show that most strategies have  $p_c(X) \approx 0$  or  $f_{\text{eq}}^\Delta(X) \approx 0$ , see Fig. 2A and B. Intuitively, in order to have a high self-cooperation rate, the strategy must be generous enough to maintain cooperation in the presence of errors. But to reach a high abundance, the strategy must be strict enough to prevent invasion by ALLC and ALLD. Within the second-order strategies, no strategy can solve this trade-off (Fig. 2A). Even L3 and L6, the two leading-eight strategies in  $\mathcal{S}_2$ , do not perform well. While L3 has a high self-cooperation rate, it is rarely adopted when competing with ALLC and ALLD. L6 is weak with respect to both dimensions.

L3's low abundance may come as a surprise, because this strategy has recently been shown to be evolutionarily stable (56, 57). However, evolutionary stability only takes into account how a strategy performs when common. For a more comprehensive view of a strategy's performance, it is also important to consider how likely that strategy can invade when rare. L3 is bad at invading ALLD. The respective fixation probability  $\rho_{\text{ALLD} \rightarrow \text{L3}}$  is close to zero. Moreover, the dynamics between L3 and ALLC is almost neutral  $\rho_{\text{L3} \rightarrow \text{ALLC}} \approx 1/N$ . Consistent with previous simulation studies (46, 48), we thus find that L3 does not evolve, despite being evolutionarily stable.

Within the third-order strategies, L1, L2, and L7 have both a high self-cooperation rate and a high abundance (Fig. 2B). Here, the good performance of L1 may appear counterintuitive. After all, L1 is almost identical to L3 (Table 1). They mostly disagree on how to evaluate a bad donor who defects against a bad recipient. While L1 would assign a bad reputation, L3 assigns a good one. This difference affects how successful the two strategies are at invading ALLD (Fig. 2C and D). When L3 competes with ALLD, a substantial fraction of defectors is regarded as good (see *Inset* of Fig. 2E). These good reputations arise when bad ALLD donors defect against other bad ALLD recipients. Because those donors then gain a good reputation, ALLD gets a moderate average payoff (Fig. 2E). In contrast, when L1 competes with ALLD, defectors are systematically evaluated as bad. As a result, ALLD's average payoff is always close to the theoretical minimum (Fig. 2F). This result resonates with previous work (48) arguing that L1, L2, L7, and L8 are most successful among the leading-eight because they all have  $R(B, B, D) = B$ .

**Evolutionary Dynamics for Complete Strategy Spaces.** Next, we explore the evolutionary dynamics when all strategies of a given complexity class compete. As one may expect from the previous results, we find that among the second-order strategies, cooperation does not evolve (Fig. 3A). Here, the most frequent strategies are various variants of ALLD. Strategies with a high self-cooperation level, such as L3, are rare (Fig. 3C). For the space of third-order strategies, there is slightly more cooperation, but average cooperation levels are still below 20% (Fig. 3B). Here, we find several cooperative strategies among the most abundant strategies, such as L1 to L5, L7, and related variants (Fig. 3D).





**Fig. 2.** Evolutionary dynamics when only three strategies compete. In line with some of the previous literature (45–51, 53, 54), we first explore the dynamics when a focal strategy competes with ALLC and ALLD. For the focal strategy, we consider all second-order norms ( $S_2$ , *Top row*), and all third-order norms ( $S_3$ , *Bottom row*). (A and B) For each  $X$ , we first compute the strategy's self-cooperation level  $p_c(X)$ . This value reflects how cooperative a homogeneous population with that strategy is. In addition, we compute the strategy's abundance  $f_{eq}^A(X)$  according to the evolutionary process with ALLC and ALLD. For a strategy to maintain cooperation, both values should be close to one. (C and D) We also record which strategies are most likely to invade in a resident population of defectors. Those having the highest fixation probabilities are shown in descending order. The color indicates the self-cooperation level of the mutant strategies. The dashed line indicates the neutral fixation probability of  $1/N$ . As indicated by the figure, no self-cooperative strategies can invade ALLD within  $S_2$ . Some self-cooperative strategies can invade ALLD within  $S_3$ . The parameters used here are  $b = 5$  and  $\epsilon = 0.01$ . (E and F) Payoffs of two different leading-eight strategies, when competing with ALLD. The total population size is  $N = 50$  and the horizontal axis indicates the number of ALLD players. The *Inset* shows a snapshot of the image matrices  $M = (m_{ij})$  from the viewpoint of the leading-eight strategy. Here, half of the population are leading-eight players and the other half are ALLD players. White and black dots indicate good and bad reputations, respectively. Parameters are  $b = 5$  and  $\epsilon = 0.02$ .

However, if we look at the same figure more broadly, as shown in the *Inset*, ALLD and other noncooperative strategies are still adopted for a substantial amount of time.

Interestingly, the set of most frequent strategies includes L3 to L5, which have not been successful in the three-strategy system (Fig. 2B). In fact, none of these strategies is good at invading ALLD (Fig. 2D). Once other third-order strategies are available, these strategies can be reached through alternative evolutionary pathways. L3 and L4 can selectively invade L2 residents. In addition, they can invade L1 and L7 through (almost) neutral drift (Fig. 2E and F). Similar pathways exist for L5. Therefore, these strategies can evolve once other leading-eight strategies are present. This pairwise invasion analysis also explains why cooperation is rare overall (Fig. 3G). After ALLD is replaced by strategies like L1, L2, and L7, these strategies are—directly or indirectly—susceptible to (almost) neutral invasion by ALLC. Compared to the leading-eight, ALLC is more robust with respect to errors. Once ALLC is common, it is easily replaced by ALLD. These new evolutionary pathways can render cooperation unstable in the complete space of third-order strategies, even though some members of  $S_3$  are successful in three-strategy competitions.

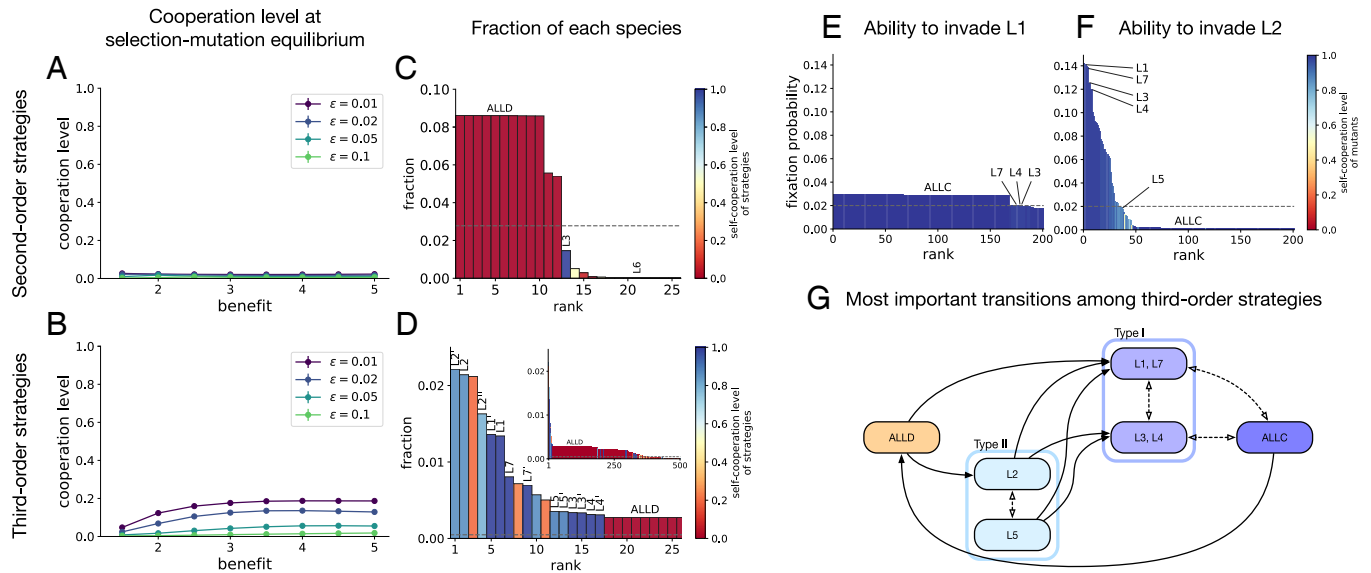
We can make sense of these evolutionary transitions by studying the similarities between strategies. The leading-eight can be classified into three types (53): Type I (L1, L3, L4, L7), Type II (L2, L5), and Type III (L6, L8). This classification is also reflected in the pathways in Fig. 3G. Strategies of different types differ in how quickly they can resolve disagreements (46). This affects how good they are at maintaining cooperation in the presence of errors, and how easily they resist invasion by ALLC.

Type I strategies have  $R(G, B, C) = G$ . This allows them to effectively contain disagreements (46). To see why, suppose everyone agrees that the donor is good, but players disagree about the recipient's reputation. In that case, observers agree that the cooperating donor should keep a good image, irrespective of their differing views of the recipient. In this way, Type I strategies prevent disagreements from spreading. As a result, they have the highest self-cooperation levels, but they are also susceptible to neutral invasion by ALLC.

Type II strategies have  $R(G, B, C) = B$ , which makes them more susceptible to disagreements caused by errors. On the other hand, they have  $R(B, B, C) = G$ , which helps to resolve disagreements. When some observers consider both the donor and the recipient as good, while others consider both to be bad, they all evaluate the cooperating donor as good. Since this event is less likely to occur than disagreements about the recipient only, this recovery path is somewhat less effective. Overall, Type II strategies still have fairly high self-cooperation levels, and they are good at resisting invasion by ALLC. However, they are easily invaded by Type I strategies.

Type III consists of the remaining leading-eight strategies L6 and L8. Under private assessment, both of these strategies are highly susceptible to errors (39, 46). Hence they cannot maintain cooperation, as also illustrated in Fig. 2B.

**Evolutionary Dynamics in Group-Structured Populations.** The previous results suggest that indirect reciprocity is unlikely to evolve in well-mixed populations. However, natural populations are rarely well mixed. They have some intrinsic structure.



**Fig. 3.** Evolutionary dynamics when all strategies of similar complexity compete. As possible strategy spaces, we again consider all second-order strategies ( $S_2$ , Top row) and all third-order strategies ( $S_3$ , Bottom row). (A and B) We display the evolving average cooperation level for different benefits of cooperation and different error rates. Error bars are not shown since they are smaller than the symbol size. The cooperation levels in (A) are almost zero for all  $\epsilon$ . (C and D) Most frequent strategies according to the stationary distribution of the evolutionary process. Colors indicate the self-cooperation level of the strategies. Names of the leading-eight ( $L1 \dots L8$ ) and their variants ( $L1' \dots L8'$ ) are shown just above the bars. Here, we defined a variant of the leading-eight as a strategy that has the same assessment rule but a different action at  $P(B, B)$ . In the *Inset* of (D), a broader range of strategies is shown. (E and F) Here, we show the strategies that are most likely to invade either a L1 or a L2 resident population. The color indicates the self-cooperation level of the mutant strategies  $p_c(X)$ . (G) A schematic diagram of the transitions between the major strategies in  $S_3$ . Solid arrows indicate transitions with fixation probabilities  $\rho$  significantly larger than  $1/N$  (here we use  $\rho > 2/N$  as the threshold). Dashed arrows indicate approximately neutral transitions ( $1/(4N) < \rho < 2/N$ ). These values are shown in more detail in *SI Appendix, Table S2*. The parameters are  $b = 5$  and  $\epsilon = 0.01$ .

Individuals are more likely to interact within their own communities. To explore whether cooperation can, in principle, evolve in structured populations, we consider a population divided into  $G$  groups. For a specific model, we use the setup of Hauert and Imhof (62, 63), which has previously been used to study direct reciprocity (64, 65). Each group consists of  $N = 50$  players (we obtain similar results for other group sizes, see *SI Appendix*). Players have all their games within their group. There, the reputation dynamics is the same as before. For the evolutionary dynamics, we assume that most of the time, players update their strategies by imitating other group members. Only occasionally do they engage in out-group imitation or adopt new strategies by mutation, as shown in Fig. 4A. As a result, groups are typically homogeneous. However, different groups may adopt different strategies. As a result, strategies of more cooperative groups are more likely to be adopted during out-group imitation. As we increase the number of groups  $G \rightarrow \infty$ , we can describe the resulting dynamics with a differential equation, see *Materials and Methods*. We explore this equation for different benefit values  $b$  and different error rates  $\epsilon$ .

Fig. 4B shows the results when players choose among all second-order strategies. Compared to well-mixed populations, we see more cooperation, even though cooperation levels still tend to be below 50%. Looking at a typical time trajectory, we observe that most groups adopt ALLD; only a minority of groups use one of the leading-eight (Fig. 4D).

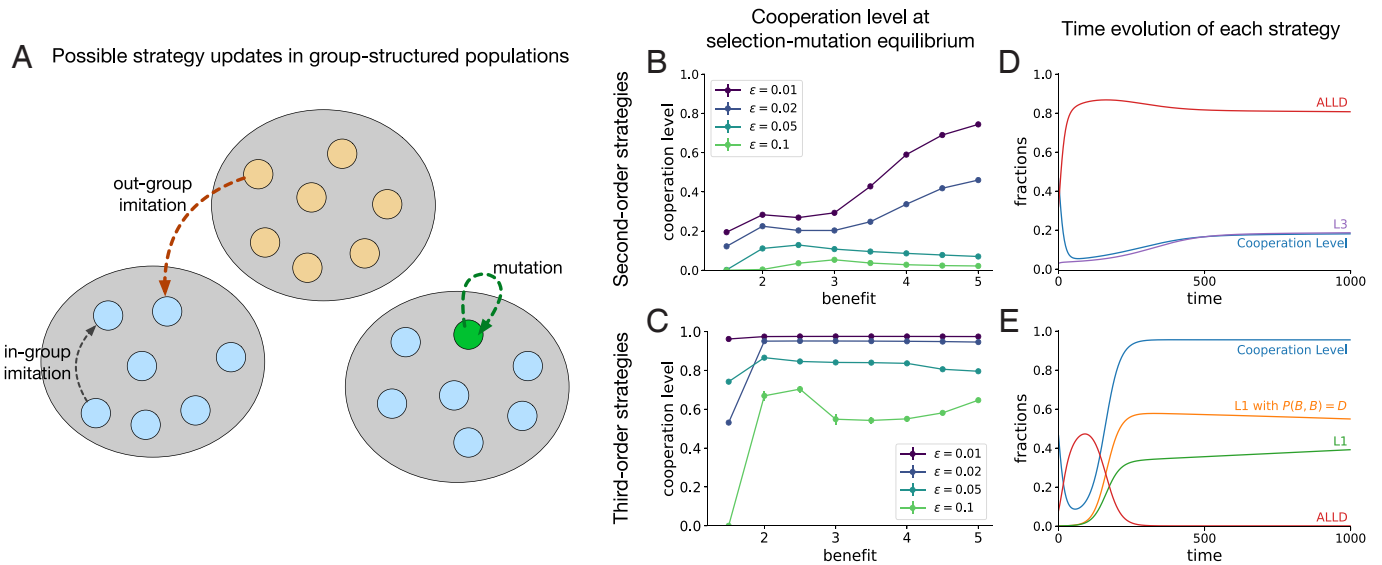
The dynamics, however, change dramatically as we allow for all third-order strategies. Here, almost full cooperation evolves for a wide range of parameters (Fig. 4C). As we look at a typical trajectory, we often observe ALLD to dominate in the beginning (Fig. 4E). But over time, more and more groups switch to L1 and to a variant of L1 (this variant differs from L1 by having  $P(B, B) = D$ ). Later, depending on parameters, also other Type I strategies such as L3 and L7 may increase in frequency. The

competition among different Type I strategies occurs at a much longer time scale, because these strategies are almost neutral. On the other hand, Type II strategies are rare, because they are selectively invaded by Type I strategies. These results suggest that within our framework of structured populations, access to third-order strategies is key for the evolution of cooperation (see also *SI Appendix, Figs. S6 and S7* for results for other parameters).

In contrast to the case of well-mixed populations, groups of Type I strategies no longer give rise to the evolution of ALLC. These neutral invasions are now prevented because there is still sufficient variety among the groups in the population. In particular, ALLD no longer goes completely extinct. Rather a small fraction of ALLD groups remain, which prevent ALLC from overtaking the entire population. This coexistence of different groups is crucial; as a result, the number of groups  $G$  needs to be sufficiently large. To confirm this intuition, we conduct Monte Carlo simulations for finite values of  $G$ . As  $G$  increases, there is a transition from a noncooperative state to a cooperative one (*SI Appendix, Fig. S8*). We observe similar results as we vary other key parameters of the model, such as the group size, or how likely errors occur (*SI Appendix, Fig. S3*). Overall, these results suggest that heterogeneity can help to stabilize cooperation, similar to previous models of direct reciprocity (65, 66).

## Discussion

Direct and indirect reciprocity are both key explanations for cooperation among nonkin (1, 2). Direct reciprocity describes how people cooperate in stable pairs. Indirect reciprocity extends the respective principle of conditional cooperation to a population level (3). While the two kinds of reciprocity follow a similar logic, respective models vastly differ in their computational complexity. Because direct reciprocity unfolds among two players



**Fig. 4.** Evolutionary dynamics in group-structured populations. (A) To complement our previous model in well-mixed populations, here we consider a setup in which individuals interact in their own subgroups. Players update their strategies by in-group imitations, out-group imitations, and mutations. Here, we assume that in-group imitations are far more frequent than the other two events. As a result, each group is homogeneous most of the time. However, different groups may have different strategies. (B and C) Cooperation levels at the stationary distribution of the evolutionary process for second-order and third-order strategies. Error bars are shown although they are smaller than the symbol size for most data points. We note that especially for higher error rates, cooperation levels can vary nonmonotonically as a function of the benefit. *SI Appendix, Figs. S6 and S7* illustrate the evolving strategies in more detail, but the effect does not seem to be driven by any particular strategy. (D and E) Here, we show a typical time series for  $b = 3$  and  $\epsilon = 0.02$ , to illustrate which strategies are most likely to emerge.

only, the game dynamics can often be described by a simple Markov chain (14). This allows for an explicit calculation of the players' payoffs. The resulting dynamics of direct reciprocity can be simulated easily even when the number of strategies is large (64–67).

For indirect reciprocity with private assessment, the situation is different. Here, the game takes place among an entire population of players. Corresponding models need to keep track of who thinks what about whom at each point in time. This makes it more difficult to compute payoffs. Results can still be derived with certain approximations (51–57) or simulations [e.g. see *SI Appendix* for more details 46–48]. Typically, however, an evolutionary approach is only feasible when the number of strategies is small. While the respective mechanistic models with a few representative strategies can be very insightful, they may also be misleading. For example, a social norm's performance against unconditional cooperators and defectors (46–48) may only give an incomplete picture of the norm's overall performance. Similarly, a norm's performance when it is common (57) may not indicate whether the norm can increase in frequency when it is rare. In some cases, it can also be difficult to determine beforehand which set of norms can be considered representative in the first place. By manually selecting a specific set of norms, we risk overlooking norms that would play an important role if they only were permitted. Here is where it becomes useful to take a systematic approach that involves all strategies of a given complexity class.

In our study, we consider the class of all second-order and third-order strategies. Among those, we find that cooperation rarely evolves when populations are well mixed. This finding may be surprising. After all, a recent study suggests that already among the second-order strategies, evolutionarily stable cooperation is possible (57). Although the respective strategy L3 is robust when common, we find that it cannot emerge when rare (Fig. 2). In addition, those strategies that are good at replacing defectors are often vulnerable to invasion by other strategies, such as ALLC (Fig. 3G).

Despite this negative result, we find that indirect reciprocity is still feasible. To make this point, we consider a particular setup of a group-structured population that has been studied earlier (62–65). In such a structured population, cooperation evolves when strategies are allowed to be sufficiently complex (Fig. 4C). When these conditions are met, cooperation is often sustained by the leading-eight strategy L1. Although L1 has been found to be effective in a previous study based on the replicator equation (9), it has not received much attention since. Among the leading-eight, L1 is the most context-independent. Cooperation is always considered as good. Defection is typically considered as bad, except when it represents an instance of “justified punishment” (where a good individual defects against a bad one). This context independence makes it easier for group members to reach a consensus. It diminishes the negative effects of disagreements that private assessment models are susceptible to (40, 41, 46). While our simulations suggest that L1 performs well in structured populations, it would be desirable to have a more analytical understanding of its properties. A natural starting point is to describe the strategy's (local) robustness against invasions, as done in previous work (52, 53, 57).

The group structure we consider is different from the one in a recent study by Kessinger et al. (51). In our study, we assume that games are only played within each group. The population's group structure only affects how strategies propagate. In contrast, Kessinger et al. (51) assume that games can take place among players from different groups. There, population structure is meant to capture the formation of “gossip groups.” Within these groups, opinions are fully synchronized. In this way, their model essentially becomes one with public assessments. As a result, they find that evolution favors the strategy L6 (Stern Judging), which we found to be ineffective. Generally, Stern Judging is only successful when opinions are perfectly synchronized (21, 22, 51, 68). Once players disagree about how they assess third parties, Stern Judging tends to amplify these disagreements. Often, this can lead to a complete breakdown of cooperation (40, 46, 57). Combined, these results suggest that indirect reciprocity either

requires a high degree of synchronization, which could be achieved through exchanging gossip (51, 52, 69), empathy (49), or institutions (50). Or it requires norms that are particularly robust with respect to disagreements, such as the norm L1 highlighted in our work. Most crucially, the two norms L1 and L6 differ in how people think of good donors who cooperate with bad recipients. Thus, these situations would serve as natural test cases for empirical work to distinguish between the two norms. One recent experiment finds that individuals are typically evaluated as good when they cooperate with bad group members (70). However, there seems to be conflicting evidence when similar experiments are performed with infants (71). This again suggests that the stability of any given social norm might be context-dependent.

More broadly, our study emphasizes the value of large-scale computational studies for evolutionary game theory. In our opinion, such studies naturally complement the strengths of analytical approaches based on simpler mechanistic models. Analytical approaches are important to form intuitions. Often times, these intuitions are more valuable than the actual numerical results, because the intuitions may also apply to other scenarios that have not been modeled explicitly. Computational methods, on the other hand, can test these intuitions in more complex environments. They can also reveal unexpected relationships that can be studied analytically later on. In our view, it takes both approaches to effectively advance the field.

## Materials and Methods

**Model of Indirect Reciprocity.** We consider a well-mixed population of size  $N$ . The members of this population (the players) repeatedly engage in donation games, see Fig. 1 A and B. Players have their own opinions about each other, which are updated each round. The state of the system is described by the image matrix  $M$ . Its elements  $m_{ij}$  represent the opinion of player  $i$  about player  $j$ . Each round, a donor and a recipient are randomly drawn from the population. The donor then decides to either cooperate (C) or defect (D) according to her strategy. When the donor cooperates, the donor pays a cost  $c$  while the recipient gets a benefit  $b$ , with  $0 < c < b$ . When the donor defects, the payoffs of both players remain unchanged. The donor's action is observed by all population members (including the donor and the recipient) with observation probability  $q$ . They all update their opinions about the donor independently, which leads to an update of the  $i$ 'th column of the image matrix. This elementary process is repeated sufficiently many times, for changing donor-recipient pairs. In the main text, we set the observation probability to  $q = 1$ . However, our further results in [SI Appendix](#) suggest that the qualitative results remain the same as we vary  $q$ .

A strategy of player  $i$  is a combination of an action rule and an assessment rule,  $(P_i, R_i)$ . We consider third-order strategies. This means that the action rule  $P_i(X, Y)$  and the assessment rule  $R_i(X, Y, A_{ij})$  are dependent on the donor's reputation  $X$ , the recipient's reputation  $Y$ , and the donor's action  $A_{ij}$ . The action  $A_{ij} \in \{C, D\}$  of donor  $i$  against recipient  $j$  is determined by the donor's action rule,  $A_{ij} = P_i(m_{ij}, m_{ij})$ . An observer  $k$ 's updated opinion about donor  $i$  is determined by her assessment rule,  $m'_{ki} = R_k(m_{ki}, m_{kj}, A_{ij})$ . Assessments are subject to errors. Each assessment is flipped with probability  $\epsilon$ .

There are  $2^8 = 256$  different assessment rules and  $2^4 = 16$  action rules. Because  $G$  and  $B$  are merely labels for two reputational states, strategies remain unchanged by swapping these labels. By removing duplicates and taking "mirror symmetries" into account, the number of unique strategies is reduced from  $2^{12} = 4,096$  to 2,080 (23). (Since 64 strategies are mirror-symmetric to themselves, there are  $(4,096 - 64)/2 + 64 = 2,080$  distinct strategies). We denote this set of strategies as  $S_3$ . For comparison, we also consider the set of the second-order rules,  $S_2$ . These are the strategies whose action and assessment rules are independent of the donor's reputation. The number of independent strategies in  $S_2$  is  $(64 - 8)/2 + 8 = 36$ , taking mirror symmetries into account.

Among the leading-eight, L3 (Simple Standing) and L6 (Stern Judging) are the only second-order strategies. All other six are third-order.

The players' long-term payoffs are defined as the average payoffs throughout all donation games:

$$\pi_i = \frac{1}{T} \sum_{t=1}^T \pi_i^{(t)}. \quad [1]$$

Here,  $T$  is the total number of time steps and  $\pi_i^{(t)}$  is the payoff of player  $i$  at time  $t$ . To this end, we define the unit of time as  $N$  donation games so that on average, each player is involved in two interactions per unit of time, one as a donor and the other as a recipient. Because there is no closed analytic formula for the long-term payoffs, we calculate them numerically by Monte Carlo simulations. We run MC simulations for  $T_{\text{init}} + T$  steps. Here,  $T_{\text{init}}$  is the number of steps for the system to reach a stationary state, and  $T$  is the number of steps for which the long-term payoffs are measured. We use  $T_{\text{init}} = 10^4$  and  $T = 10^4$ , which are sufficient for an accurate computation of the long-term payoffs.

The self-cooperation level  $\rho_c(X)$  of a strategy  $X$  is defined as the average cooperation level in a homogeneous population in which everyone adopts strategy  $X$ . Thus, the long-term payoff of a player  $i$  in such a homogeneous population is  $\pi_i = \rho_c(X) (b - c)$ .

**Evolution in Well-Mixed Populations.** At a longer time scale, players update their strategies according to an evolutionary process. In each time step, one individual is randomly chosen. With probability  $\mu$ , the individual switches to a randomly chosen strategy, chosen uniformly from the respective strategy set ( $S_3$  or  $S_2$ ). With the remaining probability  $1 - \mu$ , the focal individual  $i$  chooses a random role model  $j$  from the population. Then  $i$  imitates  $j$ 's strategy with a probability given by the Fermi rule,

$$\phi_{i \rightarrow j}^{\text{in}} = \frac{1}{1 + \exp[-\sigma(\pi_j - \pi_i)]}. \quad [2]$$

Here,  $\pi_j$  and  $\pi_i$  are the average long-term payoffs of  $j$  and  $i$ , according to Eq. 1. In particular, the larger the payoff of  $j$  is compared to the payoff of  $i$ , the more likely  $i$  is to switch. The parameter  $\sigma$  is the selection strength. It characterizes how sensitive players are to payoff differences. When  $\sigma = 0$ , the imitation probability is  $1/2$  regardless of the payoff difference; as a result, imitation occurs at random. When  $\sigma \rightarrow \infty$ , the imitation probability is 1 if  $\pi_j > \pi_i$  and 0 if  $\pi_j < \pi_i$ . Hence, imitation becomes essentially deterministic. For the simulations in the main text, we use a default value of  $\sigma = 1$ .

We consider the low mutation rate limit. In this limit, the probability that a mutant strategy  $\mathbf{q}$  fixes in a population of residents with strategy  $\mathbf{p}$  can be computed explicitly (61),

$$\rho_{\mathbf{p} \rightarrow \mathbf{q}} = \frac{1}{1 + \sum_{l=1}^{N-1} \prod_{l'=1}^l \exp\{-\sigma[\pi_{\mathbf{q}}(l) - \pi_{\mathbf{p}}(N-l)]\}}. \quad [3]$$

Here,  $\pi_{\mathbf{p}}(N-l)$  and  $\pi_{\mathbf{q}}(l)$  are the long-term payoff of a resident and of a mutant, provided that the population has  $N-l$  residents and  $l$  mutants, respectively. Using Eq. 3, the selection-mutation equilibrium of the evolutionary process can be computed by considering a reduced Markov chain with  $K$  states, where  $K$  is the number of available strategies (59). The probability to move from state  $\mathbf{p}$  to state  $\mathbf{q}$  is  $T_{\mathbf{p} \rightarrow \mathbf{q}} = \rho_{\mathbf{p} \rightarrow \mathbf{q}} / (K - 1)$ . The stationary distribution of this Markov chain is a measure for how often each strategy is adopted in the long run, after many evolutionary transitions from one strategy to another. By the ergodic theorem, the time average of any individual trajectory approaches this stationary distribution eventually. The exact time until convergence depends on the Markov chain's mixing time (which itself depends on the model parameters). We calculate the stationary distribution as the principal eigenvector of the transition matrix, using the power iteration method.

In order to calculate the fixation probability  $\rho_{\mathbf{p} \rightarrow \mathbf{q}}$  according to Eq. 3, we need to compute the expected payoff of a resident player and a mutant player for  $l = 1, \dots, N - 1$ . Thus, for each  $l$ , we run the Monte Carlo simulations of the reputation dynamics for  $T_{\text{init}} + T$  steps. Since  $\rho_{\mathbf{p} \rightarrow \mathbf{q}}$  and  $\rho_{\mathbf{q} \rightarrow \mathbf{p}}$  can be calculated simultaneously, we run the Monte Carlo simulations for  $K(K - 1)/2$



pairs of strategies. We repeat the above process for all possible combinations of strategies. In total,  $K(K-1)(N-1)/2$  Monte Carlo simulations are conducted, each of which requires  $T_{\text{init}} + T$  steps. This process needs to be repeated for different parameters, such as  $\varepsilon$  and  $q$ . (However, results for different  $b$  and  $\sigma$  can be obtained from a single simulation. They follow from the recorded cooperation probabilities between mutants and residents).

**Evolution in Group-Structured Populations.** In addition to evolution in well-mixed populations, we also consider evolution in group-structured populations (62–65). To this end, the population is subdivided into  $G$  groups of size  $N$ . Players interact with all other group members. Thus, the calculation of payoffs within a group is the same as in the well-mixed population.

Each player updates its strategy according to the following evolutionary process. At each time step, a player  $i$  is chosen randomly as the focal player. This player is then given a chance to adapt its strategy, either by intragroup imitation (with probability  $v_{\text{in}}$ ), out-group imitation (with probability  $v_{\text{out}}$ ), or mutation (with probability  $\mu$ ), as in Fig. 4A. In particular,  $v_{\text{in}} + v_{\text{out}} + \mu = 1$ .

In the case of intragroup imitation, the focal player randomly selects a role model from her own group. The focal player switches to the role model's strategy as in the case of well-mixed populations, using Eq. 2. The case of out-group imitation follows an analogous procedure. Here, the focal player  $i$  randomly selects a role model  $j$  from a different group (with all other groups being equally likely). The focal player adopts the role model's strategy with the same probability as before,

$$\varphi_{i \rightarrow j}^{\text{out}} = \frac{1}{1 + \exp[-\sigma(\pi_j - \pi_i)]}. \quad [4]$$

Out-group imitation plays a similar role as migration in genetic models of evolution. It allows strategies to move from one group to another.

The above elementary updating process is iterated for many time steps. This gives rise to a stochastic process on the space of all population compositions. In contrast to other multilevel selection models in which one group may replace another (e.g. refs. 72 and 73), selection always operates on the individual level. For our analysis, we assume that intragroup imitations occur much more frequently than mutations and out-group imitations. Specifically, while mutations and out-group comparisons happen on a similar time scale, both are rare compared to intragroup comparison,  $v_{\text{in}} \gg v_{\text{out}}$  and  $v_{\text{in}} \gg \mu$ . In this limit, all groups can be assumed to be homogeneous. However, different groups might use different strategies, because mutations might introduce novel strategies faster than out-group imitation can lead to the establishment of a single strategy.

When the number of groups  $G$  is sufficiently large, the evolutionary dynamics can be described by an ordinary differential equation (64). Let  $x_{\mathbf{p}}$  be the fraction of groups that employ strategy  $\mathbf{p}$ . Over time, these fractions can change, either because new strategies are introduced into groups by out-group imitation (and reach fixation), or they are introduced by mutations (and reach fixation). In the limit of  $G \rightarrow \infty$ , the dynamics is described by the following equation,

$$\dot{x}_{\mathbf{p}} = (1 - r) \sum_{\mathbf{q} \neq \mathbf{p}} \alpha_{\mathbf{q} \rightarrow \mathbf{p}} x_{\mathbf{q}} x_{\mathbf{p}} + r \sum_{\mathbf{q} \neq \mathbf{p}} \frac{x_{\mathbf{q}} \rho_{\mathbf{q} \rightarrow \mathbf{p}} - x_{\mathbf{p}} \rho_{\mathbf{p} \rightarrow \mathbf{q}}}{K}. \quad [5]$$

1. A. P. Melis, D. Semmann, How is human cooperation different? *Philos. Trans. R. Soc. B* **365**, 2663–2674 (2010).
2. D. G. Rand, M. A. Nowak, Human cooperation. *Trends Cogn. Sci.* **117**, 413–425 (2012).
3. M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).
4. M. A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
5. S. Uchida, K. Sigmund, The competition of assessment rules for indirect reciprocity. *J. Theor. Biol.* **263**, 13–19 (2010).
6. E. Yoeli, M. Hoffman, D. G. Rand, M. A. Nowak, Powering up with indirect reciprocity in a large-scale field experiment. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10424–10429 (2013).
7. W. Ghang, M. A. Nowak, Indirect reciprocity with optional interactions. *J. Theor. Biol.* **365**, 1–11 (2015).
8. M. Frean, S. Marsland, Score-mediated mutual consent and indirect reciprocity. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2302107120 (2023).
9. I. Okada, Two ways to overcome the three social dilemmas of indirect reciprocity. *Sci. Rep.* **10**, 1–9 (2020).
10. S. Lee, Y. Murase, S. K. Baek, Local stability of cooperation in a continuous model of indirect reciprocity. *Sci. Rep.* **11**, 14225 (2021).
11. S. Lee, Y. Murase, S. K. Baek, A second-order stability analysis for the continuous model of indirect reciprocity. *J. Theor. Biol.* **548**, 111202 (2022).

Here,  $r \equiv \mu / (\mu + v_{\text{out}})$  is the relative mutation probability compared to out-group imitation events. In this study, we use  $r = 0.05$ . The right-hand side of Eq. 5 consists of two parts. The first sum describes changes triggered by out-group imitation. The coefficient

$$\alpha_{\mathbf{q} \rightarrow \mathbf{p}} \equiv \frac{\rho_{\mathbf{q} \rightarrow \mathbf{p}}}{1 + \exp[\sigma(\pi_{\mathbf{q}} - \pi_{\mathbf{p}})]} - \frac{\rho_{\mathbf{p} \rightarrow \mathbf{q}}}{1 + \exp[\sigma(\pi_{\mathbf{p}} - \pi_{\mathbf{q}})]} \quad [6]$$

describe the flow from strategy  $\mathbf{q}$  to strategy  $\mathbf{p}$ . Here,  $\pi_{\mathbf{p}}$  is the long-term payoffs of a player in a homogeneous group of strategy  $\mathbf{p}$ . The denominator of the first term on the right-hand side describes the likelihood that a  $\mathbf{q}$ -player switches to  $\mathbf{p}$  due to out-group imitation. The numerator describes the likelihood that subsequently,  $\mathbf{p}$  reaches fixation due to in-group imitation. The interpretation of the second term in Eq. 6 is similar; it describes the possibility that a  $\mathbf{p}$ -group makes the converse transition toward  $\mathbf{q}$ .

The second sum in Eq. 5 describes changes triggered by mutations. Here, the term  $x_{\mathbf{q}} \rho_{\mathbf{q} \rightarrow \mathbf{p}} / K$  describes the probability that a player in a  $\mathbf{q}$ -group is mutating into  $\mathbf{p}$  and subsequently takes over the group. The other term  $x_{\mathbf{p}} \rho_{\mathbf{p} \rightarrow \mathbf{q}} / K$  describes the outgoing flow from  $\mathbf{p}$ -groups due to mutations. We note that the sum  $\sum_{\mathbf{p}} x_{\mathbf{p}} = 1$  by definition. Hence the equation is defined on the  $K$ -dimensional simplex.

We solve the above  $K$ -dimensional dynamical system numerically. Starting from an initial condition where every strategy is equally distributed, we integrate the differential equations until the system reaches a stationary state. We numerically confirmed that the dynamics are asymptotically stable. After a sufficiently long time ( $\approx 3 \times 10^5$ ), the system reaches a stationary state, where the fractions of strategies are time-independent. The population-wide cooperation level is calculated as the average fraction of cooperation in the stationary state:

$$\rho_c = \sum_{\mathbf{p}} \rho_c(\mathbf{p}) x_{\mathbf{p}}. \quad [7]$$

The sum is taken over all strategies in  $\mathcal{S}_2$  or  $\mathcal{S}_3$ .

**Data, Materials, and Software Availability.** All other data are included in the manuscript and/or [SI Appendix](#). The source code for this research is available in Github (74). OACIS is used for the management of simulation results (75). The generated numerical data are available in Zenodo (76).

**ACKNOWLEDGMENTS.** Y.M. thanks the Max Planck Institute for Evolutionary Biology for its hospitality during the completion of this work. Part of the results are obtained by using the Fugaku computer at RIKEN Center for Computational Science (Proposal no. ra000002). Y.M. acknowledges support from Japan Society for the Promotion of Science (JSPS) (JSPS KAKENHI; Grant no. 21K03362, Grant no. 21KK0247, and Grant no. 23K22087). C.H. acknowledges generous funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (Starting Grant 850529: E-DIRECT), and from the Max Planck Society.

12. Y. Mun, S. K. Baek, Second-order effects of mutation in a continuous model of indirect reciprocity. *Euro. Phys. J. Special Top.* **233**, 1251 (2023).
13. I. Okada, A review of theoretical studies on indirect reciprocity. *Games* **11**, 27 (2020).
14. K. Sigmund, *The Calculus of Selfishness* (Princeton Univ. Press, Princeton, 2010).
15. F. P. Santos, J. M. Pacheco, F. C. Santos, The complexity of human cooperation under indirect reciprocity. *Philos. Trans. R. Soc. B* **376**, 20200291 (2021).
16. H. Brandt, K. Sigmund, The logic of reprobation: Assessment and action rules for indirect reciprocation. *J. Theor. Biol.* **231**, 475–486 (2004).
17. M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573 (1998).
18. C. Wedekind, M. Milinski, Cooperation through image scoring in humans. *Science* **288**, 850–852 (2000).
19. O. Leimar, P. Hammerstein, Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. B* **268**, 745–753 (2001).
20. K. Panchanathan, R. Boyd, A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *J. Theor. Biol.* **224**, 115–126 (2003).
21. J. M. Pacheco, F. C. Santos, F. A. C. Chalub, Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.* **2**, e178 (2006).

22. F. P. Santos, F. C. Santos, J. M. Pacheco, Social norm complexity and past reputations in the evolution of cooperation. *Nature* **555**, 242–245 (2018).
23. H. Ohtsuki, Y. Iwasa, How should we define goodness? - Reputation dynamics in indirect reciprocity. *J. Theor. Biol.* **231**, 107–120 (2004).
24. H. Ohtsuki, Y. Iwasa, The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* **239**, 435–444 (2006).
25. H. Ohtsuki, Y. Iwasa, Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *J. Theor. Biol.* **244**, 518–531 (2007).
26. U. Berger, Learning to cooperate via indirect reciprocity. *Games Econ. Behav.* **72**, 30–37 (2011).
27. U. Berger, A. Grüne, On the stability of cooperation under indirect reciprocity with first-order information. *Games Econ. Behav.* **98**, 19–33 (2016).
28. S. Tanabe, H. Suzuki, N. Masuda, Indirect reciprocity with trinary reputations. *J. Theor. Biol.* **317**, 338–347 (2013).
29. M. Nakamura, N. Masuda, Indirect reciprocity under incomplete observation. *PLoS Comput. Biol.* **7**, e1002113 (2011).
30. N. Masuda, Ingroup favoritism and intergroup cooperation under indirect reciprocity based on group reputation. *J. Theor. Biol.* **311**, 8–18 (2012).
31. M. Nakamura, H. Ohtsuki, Indirect reciprocity in three types of social dilemmas. *J. Theor. Biol.* **355**, 117–127 (2014).
32. H. Ohtsuki, Y. Iwasa, M. A. Nowak, Reputation effects in public and private interactions. *PLoS Comput. Biol.* **11**, e1004527 (2015).
33. D. Clark, D. Fudenberg, A. Wolitzky, Indirect reciprocity with simple records. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11344–11349 (2020).
34. F. A. Chalub, F. C. Santos, J. M. Pacheco, The evolution of norms. *J. Theor. Biol.* **241**, 233–240 (2006).
35. T. Sasaki, I. Okada, Y. Nakai, The evolution of conditional moral assessment in indirect reciprocity. *Sci. Rep.* **7**, 41870 (2017).
36. J. Xu, J. Garcia, T. Handfield, "Cooperation with bottom-up reputation dynamics" in *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, N. Agmon, M. E. Taylor, E. Elkind, M. Veloso, Eds. (International Foundation for Autonomous Agents and Multiagent Systems, 2019), pp. 269–276.
37. Y. Murase, C. Hilbe, Indirect reciprocity with stochastic and dual reputation updates. *PLoS Comput. Biol.* **19**, e1011271 (2023).
38. Y. Murase, M. Kim, S. K. Baek, Social norms in indirect reciprocity with ternary reputations. *Sci. Rep.* **12**, 455 (2022).
39. S. Uchida, T. Sasaki, Effect of assessment error and private information on stern-judging in indirect reciprocity. *Chaos Solitons Fractals* **56**, 175–180 (2013).
40. S. Uchida, Effect of private information on indirect reciprocity. *Phys. Rev. E* **82**, 036111 (2010).
41. K. Oishi, T. Shimada, N. Ito, Group formation through indirect reciprocity. *Phys. Rev. E* **87**, 030801 (2013).
42. H. Yamamoto, I. Okada, S. Uchida, T. Sasaki, A norm knockout method on indirect reciprocity to reveal indispensable norms. *Sci. Rep.* **7**, 44146 (2017).
43. M. Krellner, T. A. Han, Pleasing enhances indirect reciprocity-based cooperation under private assessment. *Artif. Life* **27**, 246–276 (2022).
44. C. Perret, M. Krellner, T. A. Han, The evolution of moral rules in a model of indirect reciprocity with private assessment. *Sci. Rep.* **11**, 23581 (2021).
45. I. Okada, T. Sasaki, Y. Nakai, Tolerant indirect reciprocity can boost social welfare through solidarity with unconditional cooperators in private monitoring. *Sci. Rep.* **7**, 1–11 (2017).
46. C. Hilbe, L. Schmid, J. Tkadlec, K. Chatterjee, M. A. Nowak, Indirect reciprocity with private, noisy, and incomplete information. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12241–12246 (2018).
47. L. Schmid, P. Shati, C. Hilbe, K. Chatterjee, The evolution of indirect reciprocity under action and assessment generosity. *Sci. Rep.* **11**, 17443 (2021).
48. L. Schmid, F. Ekbatani, C. Hilbe, K. Chatterjee, Quantitative assessment can stabilize indirect reciprocity under imperfect information. *Nat. Commun.* **14**, 2086 (2023).
49. A. L. Radzvilavicius, A. J. Stewart, J. B. Plotkin, Evolution of empathetic moral evaluation. *eLife* **8**, e44269 (2019).
50. A. L. Radzvilavicius, T. A. Kessinger, J. B. Plotkin, Adherence to public institutions that foster cooperation. *Nat. Commun.* **12**, 3567 (2021).
51. T. A. Kessinger, C. E. Tarnita, J. B. Plotkin, Evolution of norms for judging social behavior. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2219480120 (2023).
52. M. Kawakatsu, T. A. Kessinger, J. B. Plotkin, A mechanistic model of gossip, reputations, and cooperation. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2400689121 (2024).
53. Y. Fujimoto, H. Ohtsuki, Evolutionary stability of cooperation by the leading eight norms in indirect reciprocity under noisy and private assessment. *PRX Life* **2**, 023009 (2024).
54. I. Okada, T. Sasaki, Y. Nakai, A solution for private assessment in indirect reciprocity using solitary observation. *J. Theor. Biol.* **455**, 7–15 (2018).
55. L. Schmid, K. Chatterjee, C. Hilbe, M. A. Nowak, A unified framework of direct and indirect reciprocity. *Nat. Hum. Behav.* **5**, 1292 (2021).
56. Y. Fujimoto, H. Ohtsuki, Reputation structure in indirect reciprocity under noisy and private assessment. *Sci. Rep.* **12**, 10500 (2022).
57. Y. Fujimoto, H. Ohtsuki, Evolutionary stability of cooperation in indirect reciprocity under noisy and private assessment. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2300544120 (2023).
58. A. Traulsen, M. A. Nowak, J. M. Pacheco, Stochastic dynamics of invasion and fixation. *Phys. Rev. E* **74**, 011909 (2006).
59. D. Fudenberg, L. A. Imhof, Imitation processes with small mutations. *J. Econ. Theory* **131**, 251–262 (2006).
60. B. Wu, C. S. Gokhale, L. Wang, A. Traulsen, How small are small mutation rates? *J. Math. Biol.* **64**, 803–827 (2012).
61. M. A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).
62. C. Hauert, L. A. Imhof, Evolutionary games in deme structured, finite populations. *J. Theor. Biol.* **299**, 106–112 (2012).
63. C. Hauert, Y. T. Chen, L. A. Imhof, Fixation times in deme structured, finite populations with rare migration. *J. Stat. Phys.* **156**, 739–759 (2014).
64. Y. Murase, C. Hilbe, S. K. Baek, Evolution of direct reciprocity in group-structured populations. *Sci. Rep.* **12**, 18645 (2022).
65. Y. Murase, S. K. Baek, Grouping promotes both partnership and rivalry with long memory in direct reciprocity. *PLoS Comput. Biol.* **19**, e1011228 (2023).
66. J. Tkadlec, C. Hilbe, M. A. Nowak, Mutation enhances cooperation in direct reciprocity. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2221080120 (2023).
67. M. Van Veelen, J. Garcia, D. G. Rand, M. A. Nowak, Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9929–9934 (2012).
68. F. P. Santos, F. C. Santos, J. M. Pacheco, Social norms of cooperation in small-scale societies. *PLoS Comput. Biol.* **12**, e1004709 (2016).
69. X. Pan, V. Hsiao, D. S. Nau, M. Gelfand, Explaining the evolution of gossip. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2214160121 (2024).
70. H. Yamamoto, T. Suzuki, R. Umetani, Justified defection is neither justified nor unjustified in indirect reciprocity. *PLoS One* **15**, e0235137 (2020).
71. J. K. Hamlin, K. Wynn, P. Bloom, N. Mahajan, How infants and toddlers react to antisocial others. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19931–19936 (2011).
72. A. Traulsen, M. A. Nowak, Evolution of cooperation by multi-level selection. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10952–10955 (2006).
73. B. Simon, J. A. Fletcher, M. Doebeli, Towards a general theory of group selection. *Evolution* **67**, 1561–1572 (2013).
74. Y. Murase, Source code. Github. [https://github.com/yohm/sim\\_evo\\_social\\_norms](https://github.com/yohm/sim_evo_social_norms). Deposited 17 May 2024.
75. Y. Murase, T. Uchitane, N. Ito, An open-source job management framework for parameter-space exploration: OACIS. *J. Phys. Conf. Ser.* **921**, 012001 (2017).
76. Y. Murase, Simulation results. Zenodo. <https://zenodo.org/records/11337579>. Deposited 27 May 2024.