

Efferson, Charles; Bernhard, Helen; Fischbacher, Urs; Fehr, Ernst

Working Paper

Super-Additive Cooperation

CESifo Working Paper, No. 10133

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Efferson, Charles; Bernhard, Helen; Fischbacher, Urs; Fehr, Ernst (2022) : Super-Additive Cooperation, CESifo Working Paper, No. 10133, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/271777>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Super-Additive Cooperation

Charles Efferson, Helen Bernhard, Urs Fischbacher, Ernst Fehr

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Super-Additive Cooperation

Abstract

Repeated interactions provide a prominent but paradoxical hypothesis for human cooperation in one-shot interactions ¹⁻³. Intergroup competitions ⁴⁻⁷ provide a different hypothesis that is intuitively appealing but heterodox. We show that neither mechanism reliably supports the evolution of cooperation when actions vary continuously. Ambiguous reciprocity, a strategy generally ruled out in models of reciprocal altruism, completely undermines cooperation under repeated interactions, which challenges repeated interactions as a stand-alone explanation for cooperation in both repeated and one-shot settings. Intergroup competitions do not reliably support cooperation because groups tend to be similar under relevant conditions. Moreover, even if groups vary, cooperative groups may lose competitions for several reasons ⁸. Although repeated interactions and group competitions do not support cooperation by themselves, combining them often triggers powerful synergies because group competitions can stabilise cooperative strategies against the corrosive effect of ambiguous reciprocity. Evolved strategies often consist of cooperative reciprocity with ingroup partners and uncooperative reciprocity with outgroup partners. Results from a one-shot behavioural experiment in Papua New Guinea fit exactly this pattern. They thus indicate neither an evolutionary history of repeated interactions without group competition nor a history of group competition without repeated interactions. Our results are only consistent with social motives that evolved under the joint influence of both mechanisms together.

JEL-Codes: C600, C700, C900.

Keywords: evolution of cooperation, reciprocity, intergroup competition, social dilemma.

Charles Efferson
Faculty of Business and Economics
University of Lausanne / Switzerland
charles.efferson@unil.ch

Helen Bernhard
Department of Economics
University of Zurich / Switzerland
helen.bernhard@econ.uzh.ch

Urs Fischbacher
Department of Economics
University of Konstanz / Germany
urs.fischbacher@uni-konstanz.de

Ernst Fehr
Department of Economics
University of Zurich / Switzerland
ernst.fehr@econ.uzh.ch

25 November 2022: We thank seminar participants at Harvard University, the Institute for Advanced Study in Toulouse, the Santa Fe Institute, and the Universities of Amsterdam, Lausanne, Konstanz, and Zurich. Funding provided by the Swiss National Science Foundation (C.E., Nr. 100018 185417/1) and the European Research Council (E.F., “Foundation of Economic Preferences”).

Although repeated interactions may seem like a paradoxical explanation for why humans cooperate in one-shot social dilemmas, the key claim is that people do not really engage in one-shot interactions. According to this explanation, the psychology that organises human decision making has evolved to treat, given the slightest reason, any interaction with a first-time acquaintance as if it were the beginning of a long-term relationship¹. This hypothesis rests on two additional claims about social interactions in ancestral groups. First, ancestral groups were typically small and cohesive, and most relationships involved interacting repeatedly with group affiliates⁶. When an ancestral human interacted with someone in the same group, chances were the pair would interact again, and reputations were at stake. Second, any uncertainty about whether an ancestral pair would interact again^{2,3} involved a crucial asymmetry. Behaving badly and damaging one's reputation was an expensive error if the pair did interact again. Behaving well and needlessly protecting one's reputation was a cheap error if the pair did not. Selection favoured risking the cheap mistake⁹.

Thus, according to this hypothesis, contemporary humans have an evolved psychology primed to treat a one-shot interaction as if it is only the first of many interactions. Put differently, even the most superficial indication that interactions might be repeated leads people to behave as if they are beginning a long-term relationship based on reciprocity and, if all goes well, mutual cooperation. Therefore, when we observe one-shot cooperation, we are actually observing the residue of ancestral conditions¹ rather than a clear-eyed response to the explicit incentives at hand. A fundamental trigger for this ancestral psychology is shared group affiliation. When interacting with someone who is, for whatever reason, in the ingroup category, one should start nice and behave reciprocally because shared group affiliation in the ancestral past was a reliable indication that interactions were likely to repeat¹⁰⁻¹². Equivalently, when interacting with someone in an outgroup category, the ancestral psychology should remain relatively idle, and one should behave selfishly.

Competitions between groups represent a very different hypothesis, also prominent, about why humans cooperate in one-shot social dilemmas^{4,5,7,13,14}. The hypothesis is not paradoxical, but a history of controversy^{15,16} means its prominence is due to infamy as much as anything else. The key claim is that cooperation was costly for the individual but beneficial for the group. Ancestral competitions between groups ensured that selection at both levels shaped evolutionary dynamics. Selfish people enjoyed an advantage over cooperative people within groups, but groups with many cooperative people enjoyed an advantage over groups with many selfish people¹⁶. If the group-selection effect was strong enough, populations would have evolved so that people were cooperative with ingroup members and selfish with outgroup members, both of which would have helped with intergroup competitions. These behaviours rested on a parochial psychology that people retain today¹⁷. The controversy surrounding this hypothesis is due to the idea that selection within groups

and migration between groups would have quickly made all groups similar in the ancestral past, and so the group-selection effect would not have been strong enough¹⁸.

Ultimately, both hypotheses seem coherent, and they lead to overlapping predictions. If the social setting has any features that lead people to see each other as group affiliates, people should behave cooperatively. Otherwise, people should behave selfishly¹⁷. The features in question can be implicit or explicit, subtle or conspicuous, a matter of conscious awareness or not¹¹. In addition to overlapping predictions, both hypotheses also rest on beliefs about ancestral social groups that will probably remain difficult or impossible to verify. How, then, can we impartially evaluate which of the two hypotheses captures the evolutionary mechanisms responsible for one-shot cooperation?

We do so with a large and comprehensive modelling project, plus a closely related behavioural experiment based on a one-shot game in Papua New Guinea. By developing both models and an experiment with parallel designs, we recruit the complementary strengths of both methods in a way that ensures the link between theory and empiricism is direct and transparent¹⁹. A direct and transparent link means that we do not need a vague intermediate step in which we extrapolate from models based on one type of social interaction to experiments involving another type of social interaction, with potentially misleading predictions as a result²⁰. In addition, we conducted our experiment with Perepkas and Ngenikas, two horticultural groups in the Western Highlands of Papua New Guinea. The Western Highlands are an ideal place to evaluate evolutionary theories of human cooperation because the people who live there, relatively speaking, are largely beyond the reach of state institutions. Social preferences, local norms, reciprocity, and group affiliation are the main forces that govern social life. These forces were probably pervasive for much of the human evolutionary past, and so they are the primary points of contention with respect to the evolution of human cooperation. In contrast, the enforceable contracts and legal institutions of contemporary large-scale societies introduce an additional force that is recent in evolutionary terms. This can confuse the interpretation of empirical findings by confounding ancestral psychologies with incentives, norms, and expectations tied to contemporary institutions.

With our jointly theoretical and experimental approach, we show that neither hypothesis about the evolution of one-shot cooperation actually works. Neither repeated interactions alone nor intergroup competitions alone support ingroup cooperation in any meaningful way, and neither mechanism leads to ingroup and outgroup predictions consistent with the behaviour observed in Papua New Guinea. Importantly, however, even though the discussion of the two hypotheses often seems to treat them as strict alternatives, they are not. Repeated interactions within groups and competitions between groups can coincide²¹. We also show that combining the two mecha-

nisms generates strong positive interactions. The result is the evolution of cooperative reciprocity with ingroup members, which amplifies cooperation within groups, and uncooperative reciprocity with outgroup members, which erodes cooperation between groups. This mix in which all equilibrium strategies are reciprocal, but not all reciprocal strategies are cooperative, is exactly what we observed among our participants in Papua New Guinea.

Repeated interactions and intergroup competition in all combinations

Our models examine the evolution of strategies for a two-person social dilemma with a continuous action space^{22,23}. The game is a theoretical version of the social dilemma we used in Papua New Guinea, and so both our models and experiment rest on the same stage game. Moves are sequential, and thus the game is an apt description of many social dilemmas past and present. Food sharing^{24,25} and alloparental care²⁶, for example, must be sequential social dilemmas with continuous action spaces. They are not simple prisoner's dilemma games in which players simultaneously decide to defect fully or cooperate fully. The emphasis on continuous action spaces is not trivial. As we will see, intuitions honed on the analysis of reciprocal altruism in repeated prisoner's dilemma games²⁷ do not extend to settings where cooperation can vary continuously.

For the stage game (Supplementary Information § 1), each player in a pair has an endowment. The first mover transfers any amount from her endowment to the second mover, and the transfer is doubled. Conditional on the first mover's transfer, the second mover transfers any amount from her endowment to the first mover, and this transfer is also doubled. Because transfers are doubled, expected relatedness cannot explain cooperation. Given what we know about average relatedness within groups in small-scale societies²⁸, efficiency gains would have to be much higher than this for relatedness alone to be an adequate mechanism.

A one-shot interaction consists of a single stage game. Repeated interactions consist of multiple stage games, with new endowments for each interaction. An individual's strategy consists of an initial transfer and a response function. The initial transfer specifies how much the individual transfers, if first mover, for the first interaction only. For all subsequent transfers, however many, the response function specifies an individual's current transfer as a function of her partner's most recent transfer.

The **repeated interactions scenario** consists of models of populations sub-divided into 40 groups of 24 individuals each without any competition between groups. Individuals within groups

pair off randomly to play the game. Individuals only have ingroup strategies, and we consider both one-shot games and repeated interactions (Supplementary Information § 2.1.4). Strategies that evolve under repeated interactions provide predictions for the ingroup treatment in our experiment, while strategies that evolve under one-shot play provide predictions for the outgroup treatment. Importantly, we ignore uncertainty about whether a game is one-shot or repeated^{2,3}, which maximises the scope for repeated interactions to support cooperation. The repeated interactions scenario isolates the effects of repeated interactions and the reputational concerns they create from the effects of intergroup competition and more generally outgroup interactions of all sorts.

The **group competition scenario** again consists of models in sub-divided populations. In this scenario, however, groups compete, and games are always one-shot. Groups are paired within a generation (Supplementary Information § 2.1.5). Each individual plays both a one-shot social dilemma with a randomly selected ingroup partner and a one-shot social dilemma with a randomly selected outgroup partner from the paired group. After game play, paired groups compete with each other with relatively low probabilities (Supplementary Information § 2.1.7) that decrease as the groups become more similar (Supplementary Information § 2.1.5). This approach reflects the idea that paired groups assess each other, and they both avoid competing when they have trouble identifying the likely winner⁴. We can think of a competition as a violent conflict, a competition for some limited resource, or a process where the culture of one group displaces the culture of another group.

Because individuals play the social dilemma with both ingroup partners and outgroup partners, strategies for doing so are explicitly conditional on group affiliation. Our experiment involved both ingroup and outgroup pairings, and so we can directly derive ingroup and outgroup predictions for the experiment from the ingroup and outgroup strategies that evolve in the model. The opportunity to cooperate with outgroup partners in our models is different from most evolutionary models of parochialism because most models limit attention to outgroup strategies that range from defection to outright aggression²⁹. Defection is thus the most generous feasible option for an outgroup interaction. In general, the group competition scenario isolates the effects of intergroup competition from the effects of repeated interactions and associated reputational concerns within groups.

The **joint scenario** combines both repeated interactions within groups and competition between groups (Supplementary Information § 2.1.6). The joint scenario is identical to the group competition scenario with one exception; ingroup interactions are always repeated. Because the joint scenario combines repeated interactions within groups and competitions between groups into

a novel selective regime, it can potentially support the evolution of strategies that differ from those that evolve when the two mechanisms operate in isolation.

A framework for examining a wide range of conditions

Many details about human social life in the evolutionary past are unknown and likely to remain that way⁶. For example, we know little about how often people moved between ancestral groups, who exactly moved when someone did move, and perhaps most importantly how these characteristics varied across populations. Extrapolating from contemporary foragers is often the best we can do^{28,30}. These limitations emphasise the importance of examining a wide range of conditions in order to identify settings that robustly support cooperation without being acutely sensitive to the details. To develop a comprehensive framework that can do this, we systematically manipulate the following six model characteristics.

1. **The dimensionality of strategy space (all scenarios).** The number of dimensions used to specify a strategy controls which strategies can and cannot arise via mutation. We vary the number of dimensions from two to four. When a strategy is two-dimensional, it consists of an initial transfer and a second quantity controlling the slope and location of a linear response function (Supplementary Information § 1.1). Possible response functions include perfect reciprocity, escalating reciprocity, and de-escalating reciprocity. Perfect reciprocity means a focal individual's transfer is exactly the same as her partner's most recent transfer (Fig. 1). When two perfect reciprocators interact, all transfers are identical to the initial transfer of the first mover. Escalating reciprocity means the focal player increases the degree of cooperation when possible (Fig. 1a), and unconditional full cooperation is an extreme case. When two escalators interact repeatedly, they converge on full cooperation, and in this sense escalation is a cooperative form of reciprocity. De-escalating reciprocity means the focal player decreases the degree of cooperation when possible (Fig. 1b), and unconditional full defection is an extreme case. When two de-escalators interact repeatedly, they converge on full defection, and thus de-escalation is an uncooperative form of reciprocity.

In a three-dimensional strategy space, a strategy consists of an initial transfer, as well as left and right intercepts for a linear response function (Supplementary Information § 1.2). Three dimensions allow for all the strategies feasible in two dimensions, but with a number of additional possibilities. For example, three dimensions allows for a class of strategies we refer to as “ambiguous reciprocity”. Ambiguous reciprocity means the focal player has a non-negatively sloped response that escalates low transfers and de-escalates high transfers

(Fig. 1c). If an ambiguous reciprocator interacts repeatedly with a partner having any positively sloped response function, the players converge on intermediate levels of cooperation (Supplementary Information § 1.2.8).

A four-dimensional strategy space adds strategies involving a wide range of nonlinear response functions (Supplementary Information § 2.3). Some of the new possibilities include nonlinear analogues of ambiguous reciprocity (Fig. 1d). New possibilities also include nonlinear forms of reciprocity that do the opposite of ambiguous strategies by de-escalating low transfers and escalating high transfers (Fig. 1d). Such strategies punish low transfers with even lower transfers and reward high transfers with even higher transfers.

2. **Cancellation effects at the individual level (all scenarios).** When a population is subdivided into groups and some individuals remain in the groups where they were born, relatedness within groups is present. When individuals play the social dilemma with ingroup partners, this relatedness allows cooperators to channel the benefits of cooperation toward other cooperators. Relatedness within groups can support the evolution of ingroup cooperation as a result, but it does not necessarily do so. Life history details, demography, and local ecological conditions can offset the effects of related individuals playing the game together³¹. Offsetting effects of this sort are called “cancellation effects”, and we refer to them as cancellation effects at the individual level to distinguish them from the group-level cancellation effects⁸ discussed below.

Our models vary cancellation effects at the individual level by implementing one of two alternate life cycles (Supplementary Information § 2.1.2). In one case, the life cycle is birth, game play, migration, group competition when relevant, and finally individual selection within groups. Game play and individual selection are **decoupled**. Individuals play the ingroup social dilemma with partners who are on average related to some extent. When individuals later compete within the group to reproduce, they compete against a different mix of group members precisely because migration occurs after game play but before individual selection. This allows related cooperators to impose the gains from mutual cooperation as a relative advantage on others who are unrelated.

In the other case, the life cycle is birth, migration, game play, group competition when relevant, and individual selection within groups. Under this life cycle, game play and individual selection are **coupled**. Relatedness within groups ensures that cooperators are relatively likely to play with other cooperators. However, because migration occurs before game play, not after, cooperators who play together also end up competing against each other to reproduce. This cancels, to some extent, the degree to which relatedness supports the evolution of cooperation³². In our case, this cancellation effect at the individual level does not completely

offset the value of playing the social dilemma with relatives. Under both life cycles, the evolution of cooperation increases with relatedness, and thus playing the game with related partners supports the evolution of cooperation (Supplementary Figures 15-16). That said, cancellation effects at the individual level also play a role in the following precise sense. In models without group competition and hence no selection at the group level, the decoupled life cycle supports more cooperation than the coupled life cycle (Supplementary Figures 15-16).

3. **Cancellation effects at the group level (group competition and joint scenarios).** Cancellation effects can also operate at the group level⁸, and the intuition parallels that at the individual level precisely. Imagine a competition between two groups, one group composed of cooperative individuals and the other of uncooperative individuals. The cooperative group wins and replaces the losing group with a descendant group that is also relatively cooperative. If the parent and descendant groups go on to compete with two entirely different groups in the subsequent generation, both groups are relatively likely to compete against less cooperative groups and thus win their respective competitions. This maximises the extent to which the group-level benefits of cooperation support the evolution of cooperation via group selection. In contrast, if the parent and descendant groups go on to compete against each other, then two cooperative groups compete against each other, with neither enjoying a relative advantage. This cancels the effects of the group-level benefits that result from both groups having many cooperative individuals.

Apart from a recent and important exception⁸, multi-level selection models are like the former example. However, if ancestral human groups did not rove freely across the landscape in search of new competitions, which seems entirely plausible, ancestral conditions were at least somewhat like the latter example. To examine this distinction, we use a novel approach to manipulate cancellation effects at the group level (Supplementary Information § 2.1.2). The 40 groups in a population constitute a population of groups. In each generation groups are paired and have a competition with positive probability. We can interpret this setting as one in which paired groups occupy adjacent territories that place the two groups in close contact. At the beginning of each generation, $\Xi \in \{0, 20, 40\}$ groups are randomly selected to enter a pool of migrating groups that move around in space. These migrating groups are randomly redistributed to the open territories. The population of groups is well-mixed when $\Xi = 40$. In this case, groups move around a lot, and groups that win intergroup competitions are relatively unlikely to compete against their descendant groups in the subsequent generation. This minimises cancellation effects at the group level. Anchoring the opposite extreme, $\Xi = 0$, groups do not move around at all, and group-level cancellation effects are as strong

as possible.

4. **The importance of differences in aggregate resources between groups (group competition and joint scenarios).** If a group competition occurs between paired groups, the group with more resources wins with a probability more or less sensitive to the difference in total resources between the two groups. We consider four levels of sensitivity (Supplementary Information § 2.1.5) controlled by the parameter $\lambda \in \{0, 10, 25, 100\}$. If $\lambda = 0$, the group that wins is unrelated to the difference in total resources. Groups compete in this case, but outcomes are unsystematic. Group selection cannot occur, and in this sense $\lambda = 0$ is like the repeated interactions scenario. As λ values increase, the group with more resources is increasingly likely to win, and the group competition and joint scenarios are increasingly different from the repeated interactions scenario.
5. **Migration rates (all scenarios).** Either eight or 16 out of 24 individuals migrate per group per generation (m_j). These values lead to relatively high or low relatedness within groups respectively (Supplementary Information § 2.1.4-2.1.6).
6. **Initial conditions (all scenarios).** In the initial generation, we seed the population with either (i) **perfect reciprocators** who initially transfer the full endowment, (ii) unconditionally **selfish** individuals, or (iii) individuals having **random** strategies drawn from a uniform distribution over the strategy space (Supplementary Information § 2.1.8). Perfect reciprocators represent initial conditions that are favourable for the evolution of cooperation, while unconditionally selfish individuals represent initial conditions that are unfavourable.

Altogether, the three scenarios and six model characteristics yields 896 combinations. For each combination, we simulated 50 independent populations. In the main paper we focus on simulation results based on three-dimensional strategies. We occasionally discuss analytical results and simulation results based on two- and four-dimensional strategies. We do so in particular for the repeated interactions scenario, where the dimensionality of the strategy space has a remarkable influence that analytical results help to clarify (Supplementary Information § 1.2.12). The Supplementary Information includes many additional results and analyses, including those that go beyond the core project outlined here, and we mention these results below in appropriate places.

The limits of repeated interactions

As a stand-alone mechanism, repeated interactions have an inexorable weakness. Cooperative strategies only evolve and persist if we arbitrarily restrict the set of possible strategies (Fig. 2).

When using two dimensions to specify strategies, reciprocal strategies that support cooperation invade and persist under a wide range of conditions. When using three or four dimensions, such strategies often invade, but they never persist. This result does not depend on whether the migration rate is relatively low (Fig. 2a – 2b) or high (Fig. 2c – 2d), nor on a specific life cycle with associated cancellation effects at the individual level (Fig. 2a and 2c vs. 2b and 2d). Increasing the number of interactions from 100 to 1000 has little to no effect (Supplementary Figures 15 – 16). Quadrupling instead of doubling transfers also leaves cooperation at very low levels (Supplementary Figure 16). Lastly, increasing relatedness by reducing migration rates to nearly zero, unrealistic for human populations^{6,28}, only supports small increases in cooperation (Supplementary Figures 15 – 16). Importantly, the weakness of repeated interactions holds even though we limit attention to dyadic interactions in which decision makers never make mistakes, two conditions that bolster reciprocity^{6,18}.

Under repeated interactions, the key distinction is between a two-dimensional strategy space (Figs. 1a - 1b) that does not allow ambiguous reciprocity and a three-dimensional space that does (Fig. 1c). A fourth dimension (Fig. 1d) has few additional evolutionary consequences. The distinction between two and three dimensions is critical for the following reason. Regardless of dimensionality, repeated interactions often lead populations to evolve at first so that most individuals exhibit high initial transfers and escalating reciprocity. Once these strategies are common, variation among individuals in the degree of escalation (Fig. 1a) is not especially important. When initial transfers are high, players start cooperating at high levels, and they have little room to escalate. One degree of escalation is just about as good as any other in terms of the behaviours generated. Selection on the exact degree of escalation is thus relatively weak and drift correspondingly important.

Crucially, however, one degree of escalation is not just about as good as any other in terms of susceptibility to invasion. Some forms of escalation are susceptible to invasion by ambiguous strategies, while others are not. We show, in fact, that the equilibrium degree of escalation is often resistant to invasion by ambiguous strategies, but it is also extremely similar to other degrees of escalation that are susceptible to invasion (Supplementary Information § 1.2.12). This means that only a tiny amount of drift should render populations vulnerable to ambiguous strategies soon after high initial transfers and escalation become common. This vulnerability is irrelevant if we exclude ambiguous reciprocity by fiat, as we do when we force the strategy space to be two-dimensional. Otherwise, this vulnerability has a dominant effect on evolutionary dynamics (Supplementary Information § 1.2.12).

Interestingly, however, when ambiguous strategies invade, they do not persist because de-escalating strategies displace them. We find no evidence for an equilibrium in which ambiguous strategies are common, and thus we should not expect to observe much ambiguous reciprocity empirically. This is no reason, however, to exclude the strategies theoretically. To do so would assume that ambiguous strategies are somehow impossible for human social cognition to manage. If we allow ambiguous strategies, dynamics take the following stylised form. Escalating strategies invade and then drift. Ambiguous strategies then invade, only to be displaced by de-escalating strategies (Supplementary Figures 9 – 14). The result is that repeated interactions, given a minimal degree of strategic flexibility, have no meaningful effect and do not provide a robust explanation for the evolution of cooperation. Both one-shot interactions and repeated interactions lead to the evolution of low initial transfers and de-escalating reciprocity (Supplementary Information § 2.1.10 – 2.1.12). When we take the strategies that evolve under repeated interactions as a prediction for how people should play with ingroup partners, we correspondingly predict low initial transfers with de-escalation, an uncooperative form of reciprocity. When we take the strategies that evolve under one-shot interactions as a prediction for how people should play with outgroup partners, we predict the same form of uncooperative reciprocity.

The limits of group competition

As a stand-alone mechanism, intergroup competition also entails a fundamental weakness. Unlike repeated interactions, the dimensionality of strategy space does not matter (Supplementary Information § 3). In contrast, the life cycle, the importance of differences in resources between groups (λ), and migration (m_j) are all jointly critical. Intergroup competition supports the evolution of ingroup cooperation only if (i) the life cycle couples game play and individual selection, (ii) λ takes the highest value we consider (i.e. $\lambda = 100$), and (iii) the migration rate is relatively low ($m_j = 8$). Otherwise, ingroup strategies evolve to generate little or no cooperation (see GC(1) for Figs. 3 – 6).

In effect, intergroup competition can support the evolution of ingroup cooperation, but a quite delicate mix of at least three characteristics has to be in place. Migration must be sufficiently low to generate meaningful differences between groups. Intergroup competitions must also have outcomes sufficiently sensitive to the group-level differences in resources that exist. Lastly, the timing of key events in the lives of individuals must take the correct form. Recall that under repeated interactions the decoupled life cycle is relatively favourable for cooperation because it reduces cancellation effects at the individual level (Supplementary Figures 15 – 16). We find exactly the opposite pattern

for the group competition scenario. The decoupled life cycle is relatively unfavourable for cooperation because it separates a group's productivity from its ability to win intergroup competitions. Specifically, the decoupled sequence includes game play, migration, and then intergroup competition. During game play, a group with many cooperative individuals enjoys large gains because many group members cooperate. Immediately after game play, however, group members migrate, and they carry the gains from cooperation with them. This redistributes resources across groups just before the group competition stage, and all else equal this attenuates the bite of intergroup competition as a mechanism. This effect is weak if the migration rate is low and strong if high.

In sum, with a delicate three-part mix in place, intergroup cooperation supports the evolution of high initial transfers with escalation for ingroup play and low initial transfers with de-escalation for outgroup play. If we remove anything from the delicate mix, both ingroup and outgroup strategies evolve to exhibit low initial transfers with de-escalation (Supplementary Information § 2.1.13 – 2.1.14). Thus, intergroup competition does not provide a robust explanation for the evolution of cooperation, and we predict uncooperative reciprocity for both ingroup and outgroup play.

Super-additive cooperation when combining repeated interactions and group competition

In the joint scenario (Supplementary Information § 2.1.6), cooperative ingroup strategies often invade and persist under circumstances in which neither repeated interactions nor group competition support the evolution of cooperation on their own. Equivalently, repeated interactions within groups and competitions between groups often interact strongly when combined. The result is a form of super-additive cooperation that far exceeds what we obtain by summing the cooperation levels from the two constituent scenarios (Figs. 3 – 6). Fig. 6d, for example, shows that the joint scenario can support the evolution of extreme super-additive cooperation even when initial conditions are unfavourable, individual-level and group-level cancellation effects are as strong as possible, and the migration rate is high. The joint scenario does not always lead to high cooperation, but it does so in a robust way that is not hypersensitive to the details. Unlike the repeated interactions scenario, the joint scenario does not require arbitrary restrictions on the strategy space for ingroup cooperation to evolve (Supplementary Information § 3). Unlike the group competition scenario, low migration, group competitions with outcomes sensitive to between-group differences (λ), and the coupled life cycle do not constitute a delicate mix in which all the pieces must be in place.

Instead, super-additivity and high levels of ingroup cooperation evolve under a wide range of conditions (Supplementary Information § 3 – 4) that involve, among other sources of variation, (i)

three- and four-dimensional strategy spaces (Supplementary Figures 138 – 155), (ii) high migration rates (e.g. Figs. 3 – 6, panels d; Supplementary Information § 3), (iii) λ values well below the maximum (e.g. Figs. 3b and 5b), and (iv) the decoupled life cycle (e.g. Figs. 3a and 5a). Finally, cancellation effects at the group level can be quite detrimental to the evolution of ingroup cooperation, but the joint scenario leads to the evolution of ingroup cooperation far more robustly than group competition alone (Supplementary Information § 2.1.18).

Why do repeated interactions and group competition interact positively? As explained above, under repeated interactions a finite population typically drifts to regions of strategy space that are susceptible to invasion by ambiguous forms of reciprocity, with the collapse of ingroup cooperation the inevitable result. Before this collapse, however, most individuals have cooperative strategies characterised by high initial transfers and escalating reciprocity (Supplementary Information § 1.2.12 and § 2.1.12). In this sense, repeated interactions support a kind of cooperative attractor that exists but is exceedingly fragile in finite populations. In the joint scenario, we augment repeated interactions with intergroup competitions, but these competitions do not create a cooperative equilibrium out of thin air. Rather, they help stabilise a finite population of cooperative escalating reciprocators against the corrosive effects of drift and ambiguous reciprocity.

This summary does not mean that repeated interactions are the core mechanism, with group competitions playing an ancillary role. Intergroup competitions are essential. Repeated interactions without group competitions lead to the evolution of uncooperative reciprocity (Fig. 7a – 7b), just like group competitions without repeated interactions (Fig. 7c – 7d). When the two mechanisms are combined, however, the result is often super-additive. Outcomes take an entirely different form compared to when the mechanisms operate in isolation. When outcomes are super-additive, ingroup strategies evolve so that cooperative forms of reciprocity prevail (Fig. 7e – 7f). Evolved outgroup strategies, in contrast, consist of uncooperative forms of reciprocity characterised by low initial transfers and de-escalation (Supplementary Information § 2.1.17). These outgroup strategies are more cooperative than the unconditional defection that represents the upper limit of feasible generosity towards outgroup interactions in many models²⁹, but they are less cooperative than the reciprocal escalation that evolves to manage ingroup interactions in the joint scenario. The joint scenario is the only scenario that generates this ingroup-outgroup pattern with any meaningful regularity.

Sequential social dilemma with Ngenikas and Perepkas

Ngenikas and Perepkas (Supplementary Information § 5) are two groups who, at the time of the experiment, inhabited territories separated by about 30 kilometres in the Western Highlands of Papua New Guinea. Although each group was aware of the other's existence, no one had any memory of hostilities between the two groups. With adult participants, we implemented a sequential social dilemma that included both ingroup and outgroup pairings (Supplementary Information § 6 – 7).

The players in a pair were each provided with an endowment of five Papua New Guinean Kina. This endowment was roughly half of a high daily wage for a labourer in the informal sector of the local workforce. Most participants earned less than this daily wage on average because they were not working for money on a regular basis. After receiving the endowment, the first mover in a pair transferred some amount between zero and five Kina, in increments of one Kina, to the second mover. The experimenter doubled this transfer.

Before learning the amount actually transferred, the second mover specified an amount she wished to back transfer to the first mover for each of the first mover's six possible transfer levels. This is the strategy method of eliciting second mover responses, and previous research has shown it to be a reliable method for measuring behavioural strategies³³. After eliciting the second mover's strategy, the experimenter revealed the amount actually transferred by the first mover and implemented the appropriate back transfer. The experimenter also doubled the back transfer.

Using a between-subjects design, we implemented four treatments that differed in terms of the group affiliations of the two players. We varied affiliations in all combinations, which yielded two ingroup treatments (Ngenika/Ngenika and Perepka/Perepka) and two outgroup treatments (Ngenika/Perepka and Perepka/Ngenika). All players knew the rules of the game. Each player also knew the group affiliation of her partner. The experimenter mediated all interactions in private, and so interactions were anonymous apart from information about group affiliations.

As explained above, when super-additive cooperation occurs under the joint scenario, it rests on an evolved strategy profile consisting of high initial transfers with escalating reciprocity for ingroup partners and low initial transfers with de-escalating reciprocity for outgroup partners (Fig. 7 and Supplementary Information § 2.1.17). Perepkas and Ngenikas exhibited exactly this pattern when playing a sequential social dilemma equivalent to the stage game used in our models. First movers (Fig. 8a) exhibited high initial transfers with ingroup partners and low initial transfers with outgroup partners (ordinal logistic regression with ingroup dummy, $p < 0.001$). Among

second movers, response functions were positively sloped (ordinal logistic regression with standard errors clustered on subject, $p < 0.001$, see Supplementary Information § 8) and uniformly more cooperative with ingroup partners than with outgroup partners (ordinal logistic regression with standard errors clustered on subject, $p < 0.001$, see Supplementary Information § 8). Of particular importance, second movers (Fig. 8b) exhibited escalating reciprocity with ingroup partners and de-escalating reciprocity with outgroup partners³⁴. Our models show that only the joint scenario reliably predicts this strategy profile.

Discussion

Repeated interactions alone cannot explain the evolution of one-shot cooperation because they cannot explain the evolution of repeated cooperation. Without unjustifiable restrictions on the strategy space, repeated interactions always lead to the evolution of uncooperative forms of reciprocity. A few analogous results hold for the standard prisoner's dilemma in which players simply choose defect or cooperate¹⁸. One expands the strategy space in this case by allowing strategies that condition the current choice on an increasing number of past interactions. Instead of only allowing strategies that condition on the most recent interaction, with tit-for-tat a prominent example, one also allows strategies that condition on the last two interactions, the last three, or many more. Increasing strategic flexibility in this way undermines sustained cooperation^{2,3,35}, but one might object that this result requires decision makers to have unreasonably long and detailed memories. This objection does not hold for our models because we only consider strategies that require people to remember a single event.

Nonetheless, when cooperation varies continuously, strategic flexibility readily arises in other ways. If we only allow escalating and de-escalating strategies, as with our two-dimensional models, we make the continuous game similar to the standard prisoner's dilemma. Specifically, as escalating or de-escalating reciprocity become common, interacting players tend to converge quickly on full cooperation or full defection respectively. Players do not converge on intermediate levels of cooperation, which minimizes the role of continuous actions by making the continuous game similar to the standard game. Without this restriction, repeated interactions do not reliably support the evolution of cooperation.

Intergroup competitions also do not reliably support the evolution of one-shot cooperation, and we have examined key subtleties suggesting the limitations of group competition could be even more serious than imagined. In particular, group selection does not necessarily occur just

because groups compete; cooperative groups must also win^{18,36}. We have considered three reasons this may or may not happen. First, the timing of life events can affect the link between a group's productivity and its ability to win intergroup competitions. Under our coupled life cycle, a group with many cooperators produces large gains that remain in the group to help win competitions against other groups. Under our decoupled life cycle, however, migration exports the gains from cooperation before such competitions occur, which attenuates the link between productivity and winning. Migration does not hinder cooperation simply by making groups similar; migration makes groups similar at the worst possible time.

Second, cancellation effects at the group level⁸ undermine cooperation that would otherwise evolve. Intuitively, if a cooperative highly productive group ends up competing against its descendant groups, it competes against other cooperative highly productive groups. Cooperative groups enjoy little relative advantage because they compete against each other, while the uncooperative groups pair off to fight their own battles. Third, even if a productive group competes with an unproductive group, the outcome is not certain. The unproductive group may win, for example because it just happens to have a uniquely talented strategist. We manipulated the probability of such outcomes (λ) as a way of capturing all other forces that undercut the association between group competition and group selection.

Our results show that many pieces have to come together for group competition to support cooperation as a stand-alone mechanism. As we know, groups must be different from each other for the selection of groups to be meaningful^{18,36}. Equally critical, the gains from cooperation must stay in the group until group competitions occur, cooperative groups must compete specifically with uncooperative groups, and the outcomes of competitions must be sensitive to the differences between cooperative and uncooperative groups. The existence of so many necessary conditions highlights a fundamental point. Estimating the frequency and lethality of ancestral wars^{37–39}, to take a contentious example, is not by itself particularly conclusive when evaluating the role of group selection. We would also need to know which specific groups fought against each other and who exactly died when wars occurred.

Because neither repeated interactions nor intergroup competitions support the evolution of cooperation by themselves, repeated interactions merit just as much infamy as group selection. Repeated interactions may seem more palatable because the effects operate via individual selection^{1,11}, but this is irrelevant if cooperation is not the result. That said, in spite of the weaknesses of the two mechanisms when separate, each can offset the weaknesses of the other when combined. Repeated interactions create a cooperative attractor that is chronically fragile, and intergroup competitions

control this fragility in situations where they do not support a cooperative attractor of their own. This super-additive mix produces the generic pattern observed in empirical studies, namely more cooperation with ingroup partners and less cooperation with outgroup partners⁷. More precisely, super-additive outcomes predict the specific nuanced pattern we observed in Papua New Guinea. Although second movers exhibited positive reciprocity with both ingroup and outgroup partners, the differences are telling. With ingroup partners, reciprocity was escalating and thus relatively cooperative, while with outgroup partners it was de-escalating and thus relatively uncooperative. Our joint scenario, and only our joint scenario, predicts these differential forms of reciprocity under a wide variety of conditions.

The ability of the joint scenario to generate the strategies observed in Papua New Guinea suggests an important point about the evolution of cooperation. The mechanisms hypothesised to support the evolution of cooperation are rarely, if ever, mutually exclusive^{20,21,35,40}, and so any examination of a single mechanism is necessarily incomplete. A mechanism with critical limitations on its own may be critically important in interaction with other mechanisms. The challenge is that systematically examining interactions is an important but daunting task. As the number of hypothesised mechanisms increases, the number of combinations grows exponentially, and researchers have thus far ignored most of the possibilities. Disputes about the evolution of human cooperation centre largely around whether some special or even unique mechanism has shaped human evolution, with our extreme reliance on culture as a leading candidate^{5,6}. However these disputes are resolved, our results highlight the important possibility that the combination of mechanisms responsible for human cooperation can also be special or even unique.

1. Hagen, E. H. & Hammerstein, P. Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology* **69**, 339 – 348 (2006). URL <http://www.sciencedirect.com/science/article/B6WXD-4J61667-1/2/0103a9d11f1d517b3f00be4260eac9ee>.
2. Zefferman, M. R. Direct reciprocity under uncertainty does not explain one-shot cooperation, but demonstrates the benefits of a norm psychology. *Evolution and Human Behavior* **35**, 358 – 367 (2014). URL <http://www.sciencedirect.com/science/article/pii/S1090513814000518>.
3. Jagau, S. & van Veelen, M. A general evolutionary framework for the role of intuition and deliberation in cooperation. *Nature Human Behaviour* **1**, 1–6 (2017).
4. Choi, J.-K. & Bowles, S. The coevolution of parochial altruism and war. *Science* **318**, 636–640 (2007).

5. Richerson, P. *et al.* Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences* **39** (2016).
6. Boyd, R. & Richerson, P. J. Large-scale cooperation in small-scale foraging societies. *Evolutionary Anthropology: Issues, News, and Reviews* (2022).
7. De Dreu, C. K., Fariña, A., Gross, J. & Romano, A. Prosociality as a foundation for intergroup conflict. *Current Opinion in Psychology* **44**, 112–116 (2022).
8. Akdeniz, A. & van Veelen, M. The cancellation effect at the group level. *Evolution* (2020).
9. Haselton, M. G., Nettle, D. & Murray, D. R. The evolution of cognitive bias. *The handbook of evolutionary psychology* 1–20 (2015).
10. Yamagishi, T., Jin, N. & Kiyonari, T. Bounded generalized reciprocity: ingroup boasting and ingroup favoritism. *Advances in Group Processes* **16**, 161–197 (1999).
11. Haley, K. J. & Fessler, D. M. T. Nobody’s watching?: subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior* **26**, 245–256 (2005).
12. Yamagishi, T. & Mifune, N. Parochial altruism: does it explain modern human group psychology? *Current Opinion in Psychology* **7**, 39 – 43 (2016). URL <http://www.sciencedirect.com/science/article/pii/S2352250X15001931>. *Evolutionary psychology*.
13. Alger, I., Weibull, J. W. & Lehmann, L. Evolution of preferences in structured populations: genes, guns, and culture. *Journal of Economic Theory* **185**, 104951 (2020).
14. Handley, C. & Mathew, S. Human large-scale cooperation as a product of competition between cultural groups. *Nature Communications* **11**, 1–9 (2020).
15. Williams, G. C. *Adaptation and Natural Selection* (Princeton: Princeton University Press, 1966).
16. Wilson, D. S. & Wilson, E. O. Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology* **82**, 327–348 (2007).
17. Gross, J. & De Dreu, C. K. The rise and fall of cooperation through reputation and group polarization. *Nature Communications* **10**, 1–10 (2019).
18. Henrich, J. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization* **53**, 3–35 (2004).

19. Muthukrishna, M. & Henrich, J. A problem in theory. *Nature Human Behaviour* **3**, 221–229 (2019).
20. Akdeniz, A. & van Veelen, M. The evolution of morality and the role of commitment. *Evolutionary Human Sciences* **3** (2021).
21. Henrich, N. & Henrich, J. *Why Humans Cooperate: A Cultural and Evolutionary Explanation* (Oxford: Oxford University Press, 2007).
22. Wahl, L. M. & Nowak, M. A. The continuous prisoner's dilemma: I. linear reactive strategies. *Journal of Theoretical Biology* **200**, 307 – 321 (1999). URL <http://www.sciencedirect.com/science/article/pii/S002251939909967>.
23. Le, S. & Boyd, R. Evolutionary dynamics of the continuous iterated prisoner's dilemma. *Journal of Theoretical Biology* **245**, 258 – 267 (2007). URL <http://www.sciencedirect.com/science/article/pii/S0022519306004255>.
24. Gurven, M. To give and to give not: the behavioral ecology of human food transfers. *Behavioral and Brain Sciences* **27**, 543–583 (2004).
25. Kaplan, H. & Gurven, M. The natural history of human food sharing and cooperation: a review and a new multi-individual approach to the negotiation norms. In Gintis, H., Bowles, S., Boyd, R. & Fehr, E. (eds.) *Moral Sentiments and Material Interests: The Foundation of Cooperation in Economic Life*, 75–113 (Cambridge: The MIT Press, 2005).
26. Hrdy, S. B. *Mothers and Others: The Evolutionary Origins of Mutual Understanding* (Cambridge, MA: Harvard University Press, 2011).
27. Axelrod, R. M. *The Evolution of Cooperation* (New York: Basic Books, 1984).
28. Hill, K. R. *et al.* Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science* **331**, 1286–1289 (2011). URL <http://www.sciencemag.org/content/331/6022/1286.abstract>. <http://www.sciencemag.org/content/331/6022/1286.full.pdf>.
29. Rusch, H. The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141539 (2014).
30. Apicella, C. L., Marlowe, F. W., Fowler, J. H. & Christakis, N. A. Social networks and cooperation in hunter-gatherers. *Nature* **481**, 497 (2012).

31. Lehmann, L. & Rousset, F. How life history and demography promote or inhibit the evolution of helping behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 2599–2617 (2010).
32. Taylor, P. D. Altruism in viscous populations: an inclusive fitness model. *Evolutionary Ecology* **6**, 352–356 (1992).
33. Brandts, J. & Charness, G. The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics* **14**, 375–398 (2011). URL <http://dx.doi.org/10.1007/s10683-011-9272-x>.
34. A minor qualification involves second mover out-group strategies (Fig. 8b). Conditional on a first mover transfer of zero, some second movers specified transfers that were small but positive, and consequently the 95% confidence interval clustered on second mover does not quite include zero. The point at which the average response function crosses the 45-degree line, however, is very low, and so individual out-group response functions generally involved strict de-escalation. This pattern is consistent with our simulations in the sense that mutations and demographic stochasticity ensure that some agents will have strategies dictating small but positive transfers in response to transfers of zero from an out-group partner.
35. van Veelen, M., García, J., Rand, D. G. & Nowak, M. A. Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences* **109**, 9929–9934 (2012). URL <http://www.pnas.org/content/109/25/9929.abstract>. <http://www.pnas.org/content/109/25/9929.full.pdf+html>.
36. Price, G. R. Selection and Covariance. *Nature* **227**, 520–521 (1970).
37. Gat, A. *War in Human Civilization* (Oxford: Oxford University Press, 2006).
38. Bowles, S. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science* **324**, 1293–1298 (2009).
39. Fry, D. P. & Söderberg, P. Lethal aggression in mobile forager bands and implications for the origins of war. *Science* **341**, 270–273 (2013). URL <http://www.sciencemag.org/content/341/6143/270.abstract>. <http://www.sciencemag.org/content/341/6143/270.full.pdf>.
40. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).

Supplementary Information Additional results and experimental protocols available at www.github.com/cmefferson/superAdditiveCooperation.

Acknowledgements CE would like to acknowledge the support of the Swiss National Science Foundation (Grant No. 100018_185417/1).

Author Contributions CE and UF designed the simulations and coded them independently. CE analysed the data produced by the simulations and interpreted the results. CE developed the associated analytical models. HB, UF, and EF designed the experiment, and HB ran the experiment. CE analysed the data from the experiment. CE synthesised results from the simulations and experiment and wrote the paper with input from UF and EF.

Competing Interests The authors declare that they have no competing financial interests.

Data/code availability The experimental data and code for simulations are available at www.github.com/cmefferson/superAdditiveCooperation.

Correspondence Correspondence should be addressed to CE (charles.efferson@unil.ch) and EF (ernst.fehr@econ.uzh.ch).

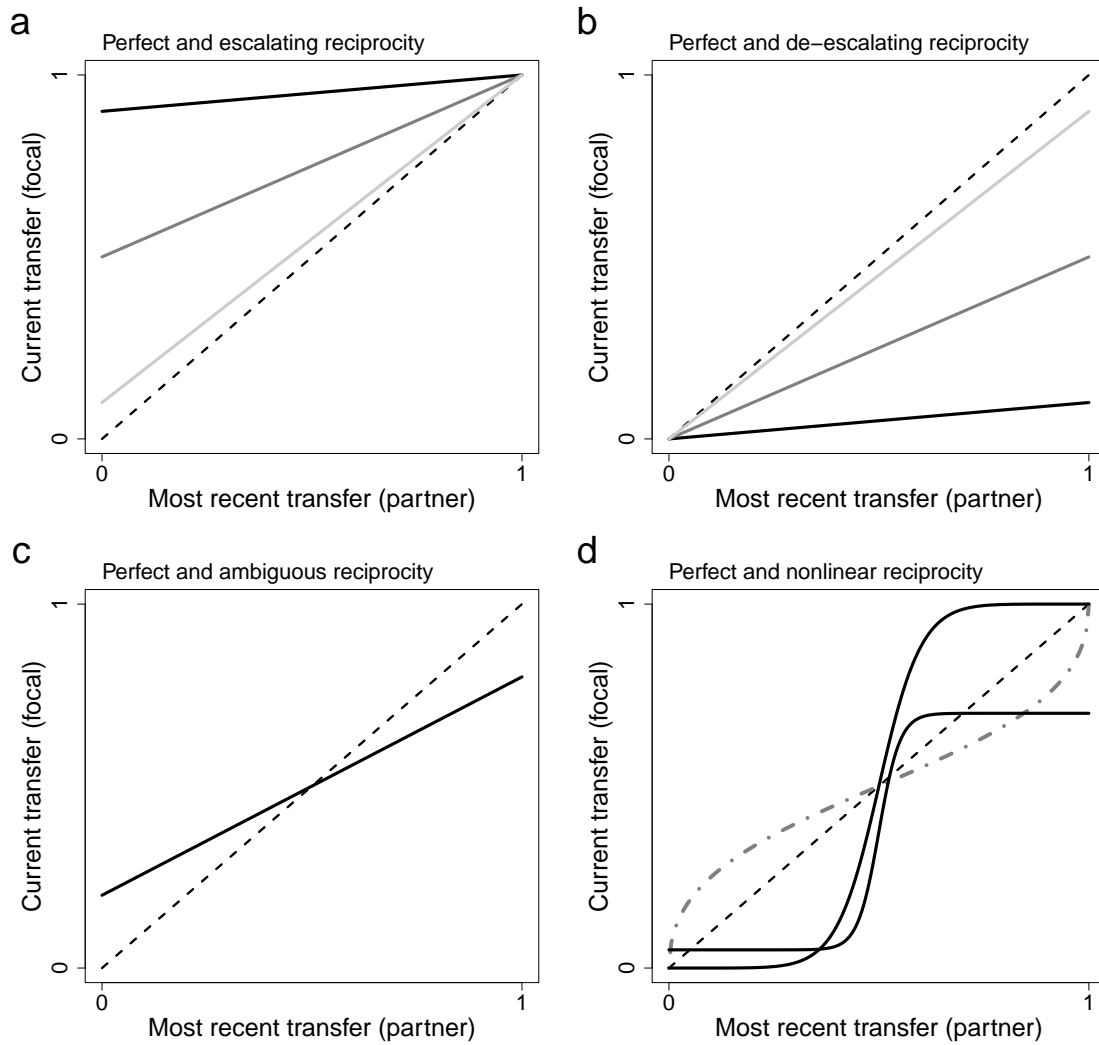


Figure 1 | Example response functions. The dashed line is the response function for perfect reciprocity. **a**, Escalating reciprocity means the focal individual has a positively sloped response function and increases the transfer level when possible. Examples include weak (light grey), intermediate (grey), and strong (black), where weak escalation increases cooperation slowly and strong does so quickly. **b**, De-escalating reciprocity means the focal individual has a positively sloped response function and decreases the transfer level when possible. Examples include weak (light grey), intermediate (grey), and strong (black), where weak de-escalation reduces cooperation slowly and strong does so quickly. **c**, Ambiguous reciprocity amounts to escalation in response to low transfers and de-escalation in response to high transfers. **d**, Nonlinear forms of reciprocity allow nonlinear analogues of ambiguous reciprocity (grey dash-dot line), as well as complex and flexible mixtures of de-escalation and escalation in response to both low and high transfers (black solid lines). Two dimensions are adequate for perfect, escalating, and de-escalating forms of reciprocity, with one dimension for the initial transfer and another for the response function. Ambiguous reciprocity requires three dimensions, and nonlinear strategies four.

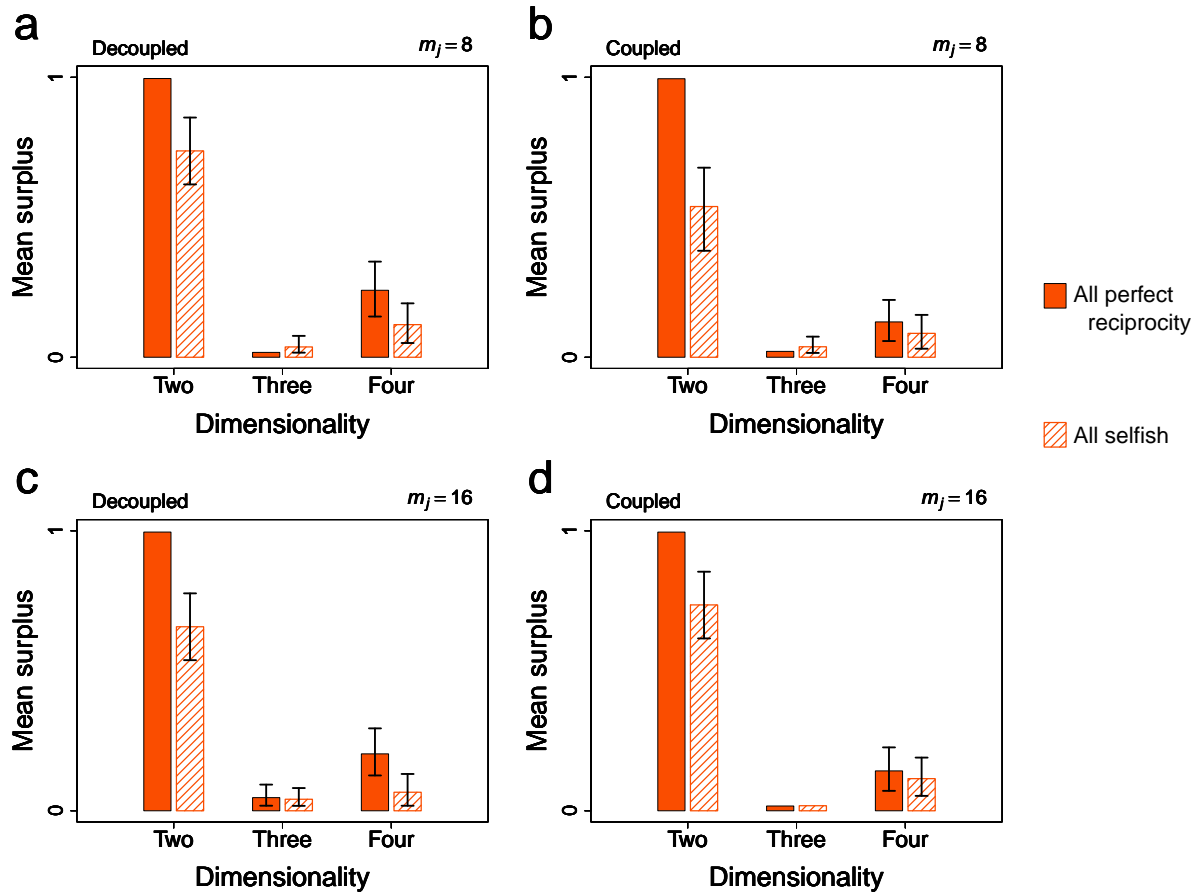


Figure 2 | Strategic flexibility hinders cooperative reciprocity under repeated interactions.

The graphs show the mean surplus per individual per ingroup interaction in the final generations of simulated evolution under $N = 100$ ingroup interactions. Error bars show 95% confidence intervals, which are calculated by bootstrapping over populations and omitted when extremely narrow. Strategies are defined in either two, three, or four dimensions (Fig. 1). Initial conditions are either relatively favourable for the evolution of cooperation (All perfect reciprocity) or relatively unfavourable (All selfish). The life cycle either decouples game play and individual selection (a, c), which reduces individual-level cancellation effects and enhances the potential for relatedness to support cooperation, or it couples them (b, d). The number of migrants per group per generation (m_j) is either relatively low (a, b) or high (c, d). Cooperative reciprocal strategies only evolve or persist when strategies are two-dimensional, and thus ambiguous reciprocity (Fig. 1c) is not possible. When strategies are defined in three or four dimensions, with a relatively high degree of strategic flexibility as a result, cooperative reciprocal strategies sometimes invade, but they never persist.

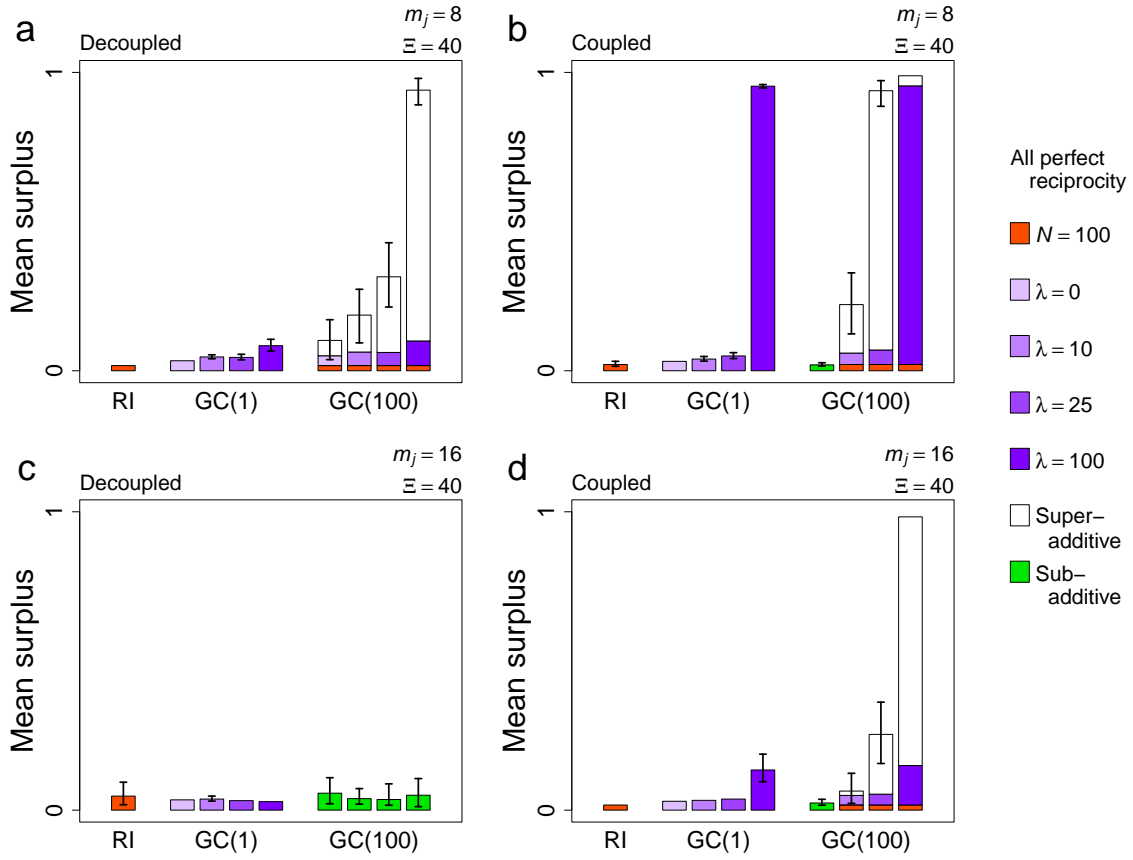


Figure 3 | Super-additive cooperation when initial conditions and group mixing are relatively favourable for the evolution of cooperation. The graphs show the mean surplus per individual per ingroup interaction from the final generations of evolutionary simulations. Error bars show 95% confidence intervals, which are calculated by bootstrapping over populations and omitted when extremely narrow. RI signifies the repeated interactions scenario with $N = 100$ interactions per ingroup pair. GC(1) indicates the group competition scenario as competition outcomes vary in sensitivity to group differences (λ). GC(100) combines the two component mechanisms into the joint scenario, with the bars from left to right corresponding to increasing λ . When the joint scenario is super-additive, the mean surplus is decomposed (Supplementary Information § 3) into the repeated interactions effect (orange), the group competition effect (purple), and the super-additive effect (white). Initial conditions consist of a population of perfect reciprocators (All perfect reciprocity), and group mixing is as high as possible ($\Xi = 40$), both of which are relatively favourable for ingroup cooperation. The life cycle either decouples (a, c) or couples (b, d) game play and individual selection. The number of migrants per group per generation (m_j) is either relatively low (a, b) or high (c, d). Panels a, b, and d show that repeated interactions and group competition interact strongly to support the evolution of ingroup cooperation when neither mechanism does so in isolation.

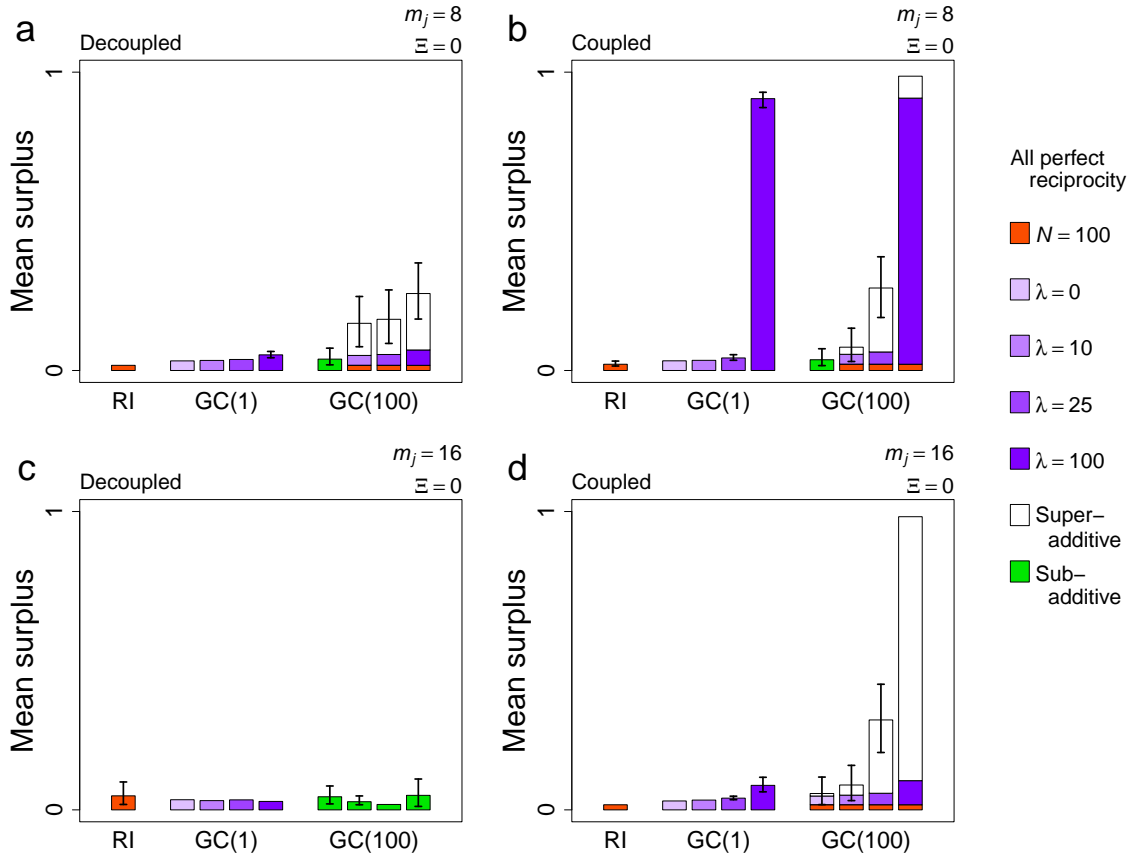


Figure 4 | Super-additive cooperation when initial conditions are relatively favourable and group mixing relatively unfavourable for the evolution of cooperation. The graphs show the mean surplus per individual per ingroup interaction from the final generations of evolutionary simulations. Error bars show 95% confidence intervals, which are calculated by bootstrapping over populations and omitted when extremely narrow. RI signifies the repeated interactions scenario with $N = 100$ interactions per ingroup pair. GC(1) indicates the group competition scenario as competition outcomes vary in sensitivity to group differences (λ). GC(100) combines the two component mechanisms into the joint scenario, with the bars from left to right corresponding to increasing λ . When the joint scenario is super-additive, the mean surplus is decomposed (Supplementary Information § 3) into the repeated interactions effect (orange), the group competition effect (purple), and the super-additive effect (white). Initial conditions consist of a population of perfect reciprocators (All perfect reciprocity), and group mixing is as low as possible ($\Xi = 0$), with the former favourable and the latter unfavourable for ingroup cooperation. The life cycle either decouples (a, c) or couples (b, d) game play and individual selection. The number of migrants per group per generation (m_j) is either relatively low (a, b) or high (c, d). Because $\Xi = 0$, cancellation effects at the group level⁸ are as strong as possible, which is generally unfavourable for cooperation. Super-additivity is nonetheless common (a, b, d), and in some cases the result of extremely strong positive interactions (d).

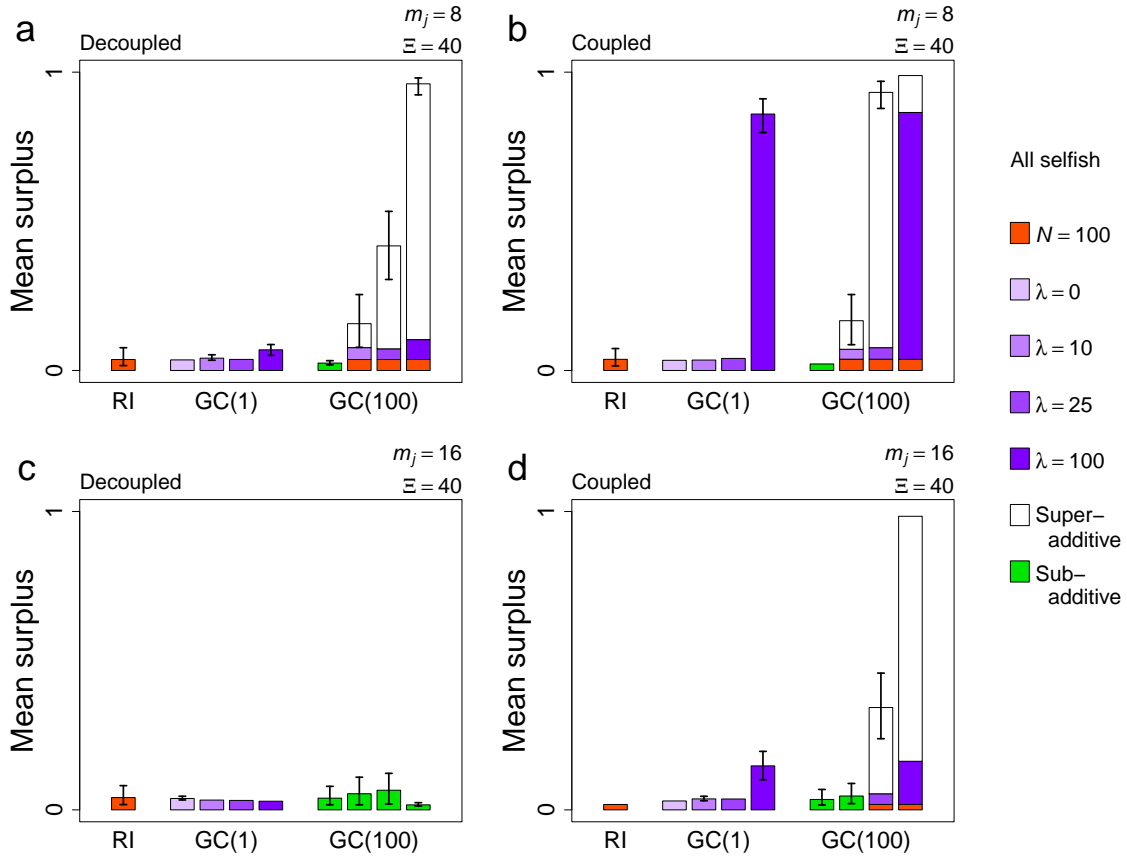


Figure 5 | Super-additive cooperation when initial conditions are relatively unfavourable and group mixing relatively favourable for the evolution of cooperation. The graphs show the mean surplus per individual per ingroup interaction from the final generations of evolutionary simulations. Error bars show 95% confidence intervals, which are calculated by bootstrapping over populations and omitted when extremely narrow. RI signifies the repeated interactions scenario with $N = 100$ interactions per ingroup pair. GC(1) indicates the group competition scenario as competition outcomes vary in sensitivity to group differences (λ). GC(100) combines the two component mechanisms into the joint scenario, with the bars from left to right corresponding to increasing λ . When the joint scenario is super-additive, the mean surplus is decomposed (Supplementary Information § 3) into the repeated interactions effect (orange), the group competition effect (purple), and the super-additive effect (white). Initial conditions consist of a population of unconditionally selfish individuals (All selfish), and group mixing is as high as possible ($\Xi = 40$), with the former unfavourable and the latter favourable for ingroup cooperation. The life cycle either decouples (a, c) or couples (b, d) game play and individual selection. The number of migrants per group per generation (m_j) is either relatively low (a, b) or high (c, d). Initially, the population consists entirely of selfish individuals, and cooperative strategies must actually invade to become established. Nonetheless, repeated interactions and group competition interact strongly to produce large super-additive gains (a, b, d).

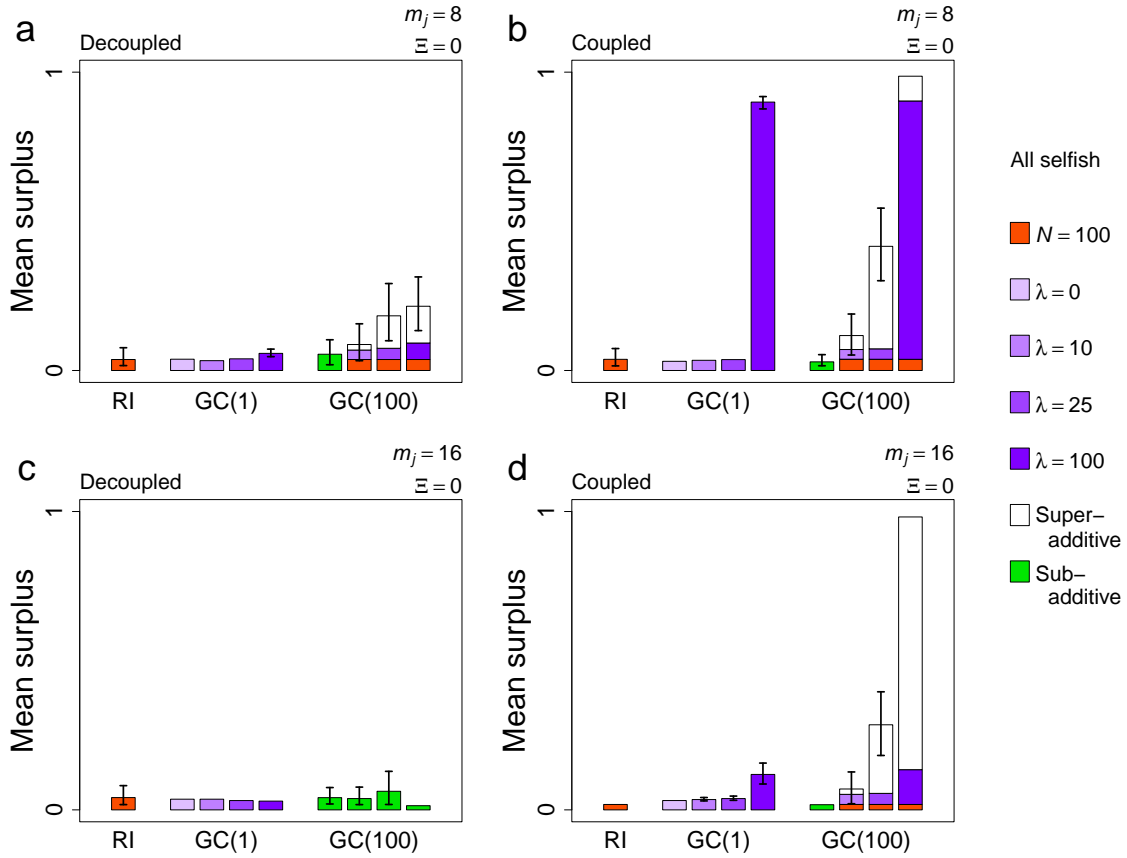


Figure 6 | Super-additive cooperation when initial conditions and group mixing are relatively unfavourable for the evolution of cooperation. The graphs show the mean surplus per individual per ingroup interaction from the final generations of evolutionary simulations. Error bars show 95% confidence intervals, which are calculated by bootstrapping over populations and omitted when extremely narrow. RI signifies the repeated interactions scenario with $N = 100$ interactions per ingroup pair. GC(1) indicates the group competition scenario as competition outcomes vary in sensitivity to group differences (λ). GC(100) combines the two component mechanisms into the joint scenario, with the bars from left to right corresponding to increasing λ . When the joint scenario is super-additive, the mean surplus is decomposed (Supplementary Information § 3) into the repeated interactions effect (orange), the group competition effect (purple), and the super-additive effect (white). Initial conditions consist of a population of unconditionally selfish individuals (All selfish), and group mixing is as low as possible ($\Xi = 0$), both of which are relatively unfavourable for ingroup cooperation. The life cycle either decouples (a, c) or couples (b, d) game play and individual selection. The number of migrants per group per generation (m_j) is either relatively low (a, b) or high (c, d). Cancellation effects at the group level⁸ hinder cooperation in general. Moreover, the initial population consists entirely of selfish individuals, and cooperative strategies must actually invade to become established. Super-additivity is nonetheless common (a, b, d), and in some cases the result of extremely strong positive interactions (d).

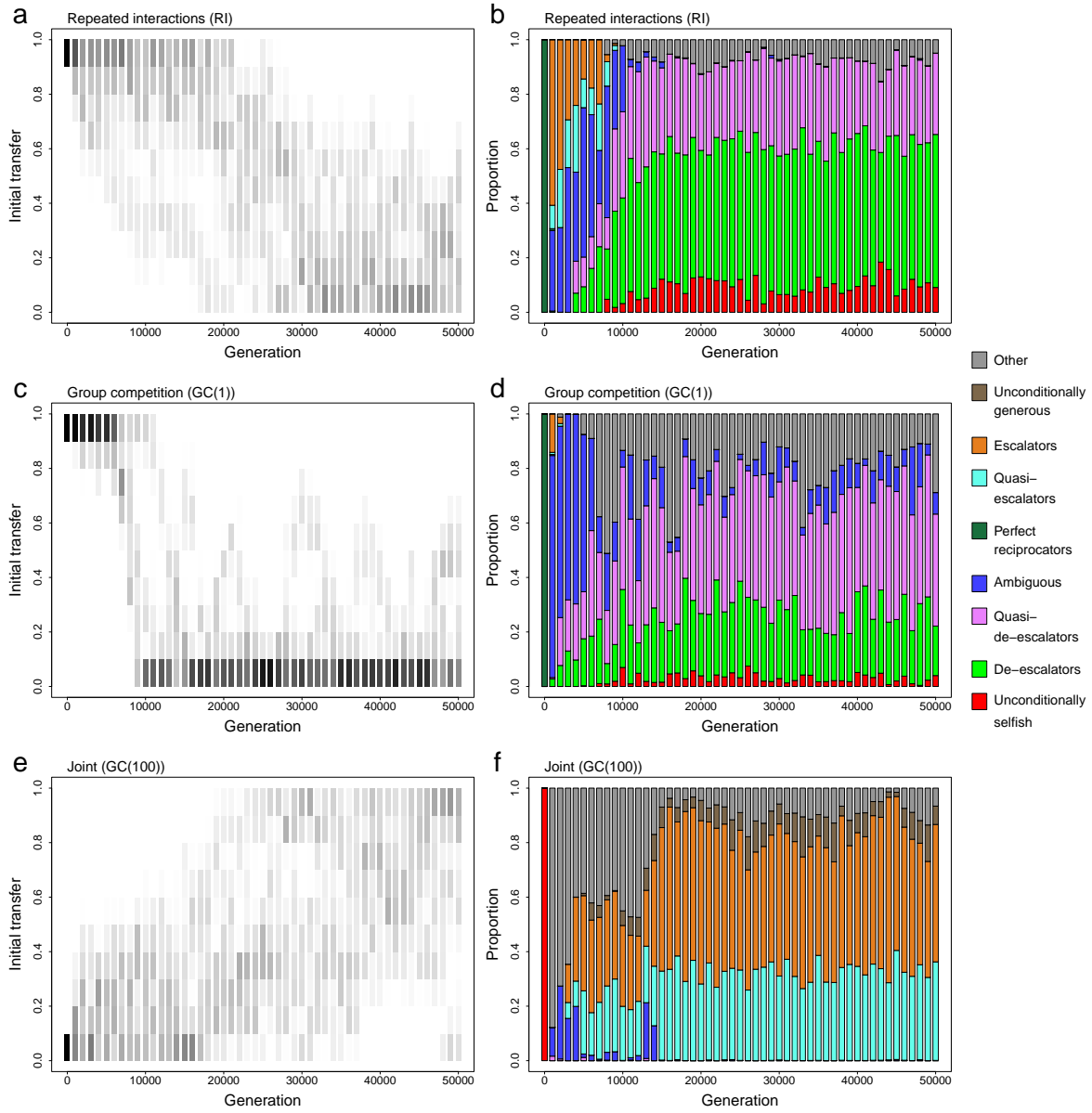


Figure 7 | Evolution of strategies under three scenarios. Panels show frequencies of initial transfers (**a**, **c**, **e**, grey scale) and response functions (**b**, **d**, **f**). We categorise response functions into discrete types. Because ambiguous strategies cover such a broad range, we separately identify extreme forms of ambiguous reciprocity (Supplementary Information § 2.1.17). Specifically, quasi-de-escalators are similar to de-escalating strategies in that they generate extremely low levels of cooperation in the long run. Similarly, quasi-escalators are similar to escalating strategies in that they generate extremely high levels of cooperation. For the results here, migration rates are high, group competitions have outcomes that are highly sensitive to differences between groups, and cancellation effects at the individual and group levels are as strong as possible (Supplementary Information § 2.1.17). Initial conditions favour cooperation in the repeated interactions (**a** – **b**) and group competition scenarios (**c** – **d**), but they disfavour cooperation in the joint scenario (**e** – **f**). Nonetheless, uncooperative forms of reciprocity prevail in the former two scenarios, while cooperative forms of reciprocity prevail in the latter scenario.

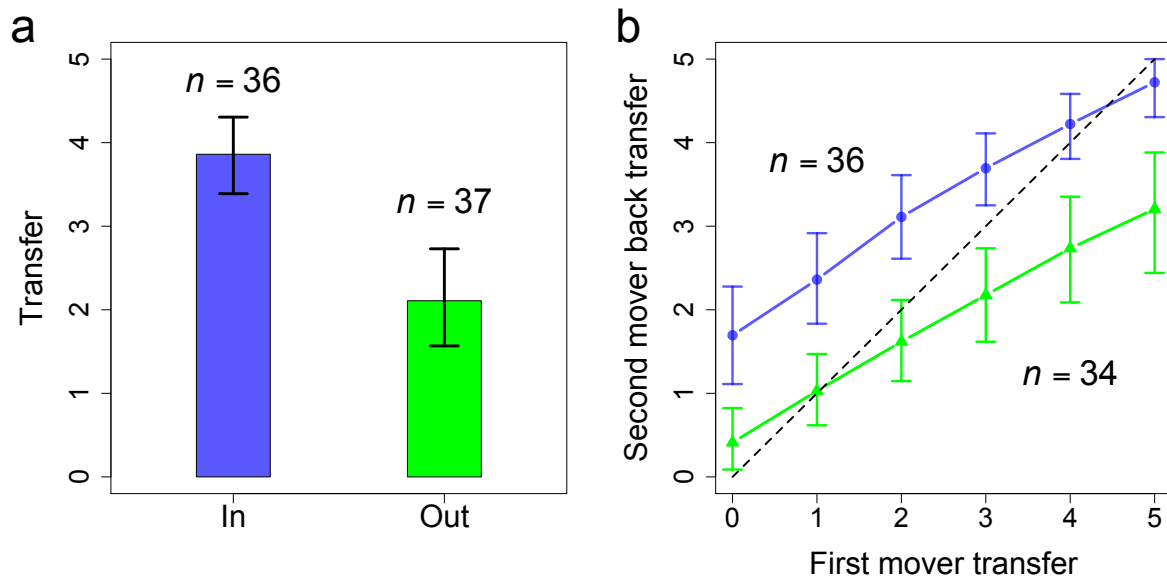


Figure 8 | First mover transfers and second mover back transfers. **a**, The mean transfers of first movers with ingroup partners (36 participants) were relatively high and with outgroup partners (37 participants) relatively low. The difference is highly significant (ordinal logistic regression with ingroup dummy, $p < 0.001$). Error bars are 95% bootstrapped confidence intervals. **b**, Second movers with ingroup partners (36 participants, blue) exhibited escalating reciprocity, and second movers with outgroup partners (34 participants, green) exhibited de-escalating reciprocity. Error bars are 95% bootstrapped confidence intervals clustered on second mover.