PAPER • OPEN ACCESS

Dynamics of heuristics selection for cooperative behaviour

To cite this article: Felipe Maciel Cardoso *et al* 2020 *New J. Phys.* **22** 123037

View the article online for updates and enhancements.

PAPER

Dynamics of heuristics selection for cooperative behaviour

Felipe Maciel Cardoso[1] , Carlos Gracia-Lázaro[1,2] and Yamir Moreno[1,2,3,*]

1. Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Spain
2. Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Spain
3. ISI Foundation, Turin, Italy
* Author to whom any correspondence should be addressed.

E-mail: yamir.moreno@gmail.com

Abstract

Situations involving cooperative behaviour are widespread among animals and humans alike. Game theory and evolutionary dynamics have provided the theoretical and computational grounds to understand what are the mechanisms that allow for such cooperation. Studies in this area usually take into consideration different behavioural strategies and investigate how they can be fixed in the population under evolving rules. However, how those strategies emerged from basic evolutionary mechanisms continues to be not fully understood. To address this issue, here we study the emergence of cooperative strategies through a model of heuristics selection based on evolutionary algorithms. In the proposed model, agents interact with other players according to a heuristic specified by their genetic code and reproduce—at a longer time scale—proportionally to their fitness. We show that the system can evolve to cooperative regimes for low mutation rates through heuristics selection while increasing the mutation decreases the level of cooperation. Our analysis of possible strategies shows that reciprocity and punishment are the main ingredients for cooperation to emerge, and the emerging heuristics would likely cooperate in one-shot interactions. Additionally, we show that if in addition to behavioural rules, genetic relatedness is included, then kinship plays a relevant role, changing emerging strategies significantly. Our results illustrate that our evolutionary heuristics model is a generic and powerful tool to study the evolution of cooperative behaviour.

1. Introduction

Game theory constitutes a powerful framework for the mathematical study of social dilemmas [1, 2]. Within this framework, the most representative and widely used game to model cooperation, the Prisoner's dilemma, has become a paradigm for modelling the evolution of cooperative behaviour [3]. The Prisoner's dilemma mimics the worst possible scenario for cooperation in which selfishness always provides a higher individual benefit than cooperative behaviour. Initial predictions indicated the social optimum would not be reachable by rational selfish individuals if the temptation for defecting (T) exceeded the reward for cooperating R. Nonetheless, cooperation is pervasive in human and animal societies [4–7], and a vast literature has demonstrated how cooperation can thrive in the presence of an appropriate evolutionary process [7–22]. The possible situations where cooperation might flourish are endless, and we are just beginning to uncover the ingredients behind the complexity observed in real systems [23, 24]. Consequently, theoretical studies usually focus on simplifications, such as individuals behaving according to fixed pure strategies [8, 14] or some arbitrary set of them [25, 26]. Yet, the reasoning and motivations of humans are more sophisticated and complex than pure strategies and decisions are usually taken factoring in many ingredients, each one with a different weight [27]. In other words, generally speaking, the selection of strategies takes place in complex systems wherein imprecise behaviour and the environment are inputs of each other in a perpetual feedback loop [28].

In this line, behavioural economics has shown that humans respond in unexpected ways [29, 30] and often seem to possess hardwired heuristics while acting in experimental situations [31, 32]. Empirical
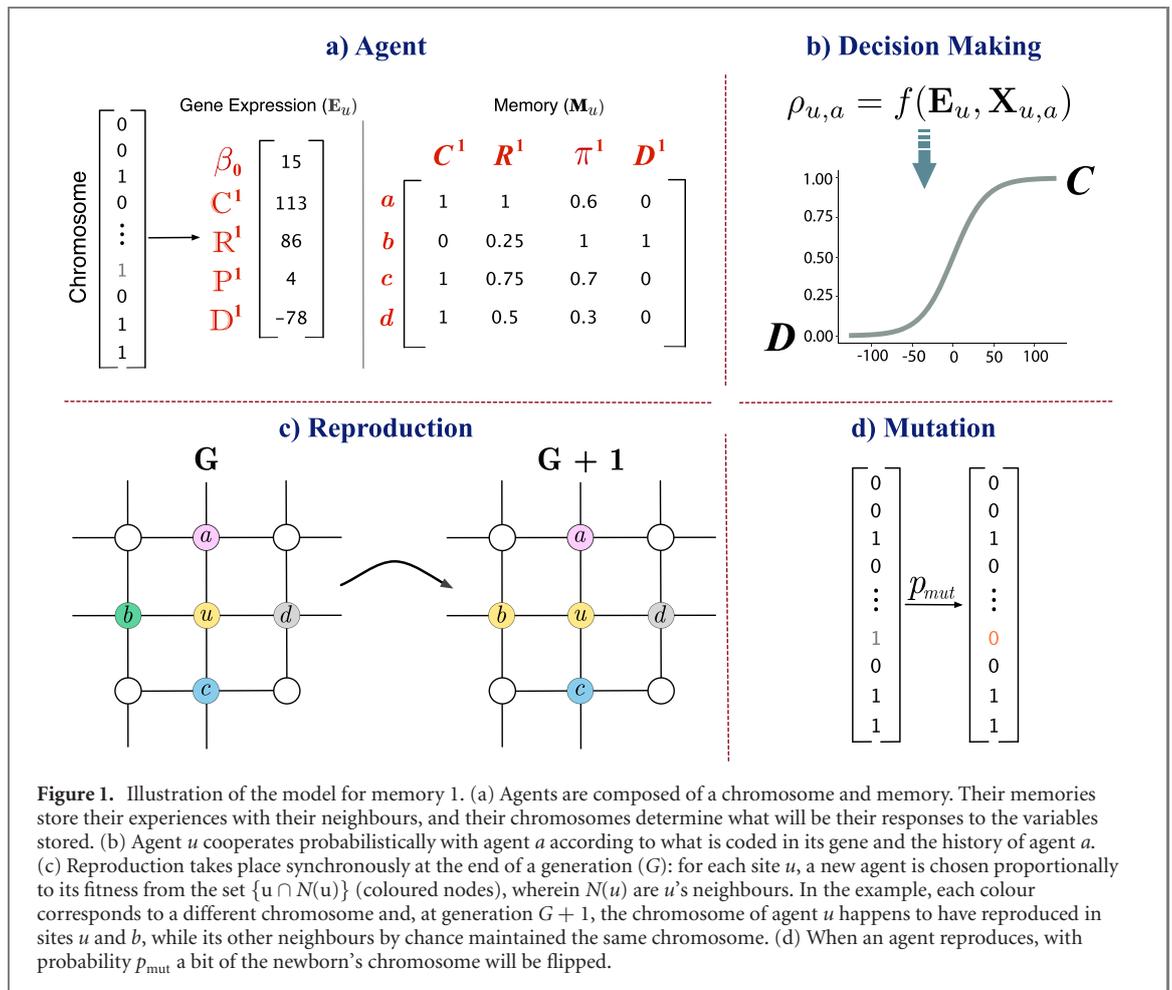
Figure 1. Illustration of the model for memory 1. (a) Agents are composed of a chromosome and memory. Their memories store their experiences with their neighbours, and their chromosomes determine what will be their responses to the variables stored. (b) Agent u cooperates probabilistically with agent a according to what is coded in its gene and the history of agent a. (c) Reproduction takes place synchronously at the end of a generation (G): for each site u, a new agent is chosen proportionally to its fitness from the set $\{u \cap N(u)\}$ (coloured nodes), wherein $N(u)$ are u's neighbours. In the example, each colour corresponds to a different chromosome and, at generation $G + 1$, the chromosome of agent u happens to have reproduced in sites u and b, while its other neighbours by chance maintained the same chromosome. (d) When an agent reproduces, with probability p_{mut} a bit of the newborn's chromosome will be flipped.

findings have also shown that humans automatic responses are modelled by experiences from daily-life, building heuristics or intuitions which tend to favour cooperation [33, 34]. Therefore, it is plausible that cooperative societies are sustained by existent heuristics, maintained by norms [35, 36] or biological factors [24, 31, 37], that have resulted from a selection dynamics. It is thus imperative to understand how such possible heuristics have evolved, which will allow explaining the ingrained mechanisms behind the behaviour observed in living beings.

In this paper, we investigate the evolution of cooperative strategies through an agent-based model of heuristics selection inspired by evolutionary algorithms [38]. The ultimate goal is to obtain a description of the evolutionary process that could lead to different strategies. Explicitly, we consider agents composed of a chromosome and memory to store information of other players' previous actions (figure 1(a)). Their actions are responses, according to what is coded in their genes, to other players' history. The strategy space is given, thus, by all the possible genes' combinations. This does not mean that we model behaviours defined by real genomes: decision making, especially in humans, has entangled layers of complexity, and such an approach would be misguided. Rather, we use chromosomes as a tool to model heuristics formed through cultural or biological evolution [28, 39].

In our framework, the fitness of agents corresponds to the payoff obtained in iterated games, and it determines the agent's reproduction rates. Offspring will inherit its parents' chromosomes while being susceptible to mutation. Note that our approach differs from elementary evolutionary algorithms: they optimize functions in a constant fitness landscape, but in evolutionary games changes in the population imply changes in the fitness landscape [40], which can be easily seen in any form of the rock-paper-scissors game [20].

The use of evolutionary algorithms to explore the adaptation of agents is not new [11, 12, 41], and previous works have studied the evolution of automata-like strategies, though aiming at answering specific situations [42, 43]. In these studies, the equivalent of a chromosome is a tool to encode an extensive set of memory-based strategies used to understand when cooperation may thrive. Unfortunately, these types of strategies are hardly realistic and do not correspond to the optimal model for understanding the mechanisms behind human or animal responses. A model of heuristics should resemble more closely

automatic responses based on intuition and past experiences [27], namely, by considering that intuitive responses are no more than stochastic processes which take as inputs the variables observed by the individual.

Here, we develop a modelling approach in which agents can evaluate different variables at the same time, thus resembling real situations wherein different factors interact and affect actions. Agents decisions are determined by an activation function taking as input their chromosome and the information to which they have access. Given their theoretical and practical importance, we focus on the evolution of cooperation in social dilemmas. For this case, therefore, we selected a set of variables based on the history of the players with whom they are playing. Nonetheless, our modelling framework is generic, and any arbitrary set of variables can be added or removed according to the question of interest. Our results show that the specified heuristics can evolve to cooperative equilibria for low mutation rates. An analysis of agents chromosome reveals that cooperation endures by reciprocity, indicating that the evolution drives heuristics to reproduce a fundamental mechanism underlying cooperation in nature, especially in humans [44, 45]. In this case, emerging strategies of conditional cooperators (CCs) dominate, permitting cooperation to prosper. Finally, we provide an extension wherein agents can evaluate their genetic relatedness with others. The population in this scenario evolves to similar equilibria. However, the agents' chromosomes differ significantly from the first model. Kin identification becomes the main mechanism of cooperative heuristics, although agents still need to have a memory of their past actions for cooperation to endure.

Undoubtedly, varied environmental or perception variables affect the resulting behaviour in humans and other animals. Unfortunately, it is not straightforward to capture which variables guided evolution to the emerged behaviour in each particular scenario. In this line, our proposal provides one generic approach for the modelling of such processes. In particular, the model here presented also contribute some insightful results with the current specifications. Namely, we observe that cooperation can spread spontaneously when memory is available, and that mutation is essential to ensure this outcome. Moreover, although the same behaviour might be observed in distinct populations, the underlying causes might be significantly different, as we observe with our kin and non-kin models. These insights suggest that our method can be a useful tool to uncover the ultimate causes behind the evolution of pro-social behaviour.

2. Methods

2.1. Population dynamics

We consider a virtual environment inhabited by n haploid agents in a zero population growth condition, each one of them (u) containing a chromosome $\mathbf{A}_u$ defining the heuristic which will guide its decision. Each agent interacts with each other through links defined by a static contact structure, in which L is the set of edges connecting the two pairs of agents. More specifically, agents are placed in regular lattices and random regular networks (RRN) (random networks of fixed degree k). In real systems, a generation embodies repeated interactions between individuals, and it is known that fast selection fluctuations can suppress cooperation even in the cases in which it is the only rational choice [18]. Thus, in our model, each generation is composed of a finite number of $s = 100$ time steps and, therefore, $s|L|$ dyadic interactions take place, i.e., one for each edge at each time step. Therefore, at each time step t, all connected agents u and v interact in a game and obtain the payoffs π_u^t and π_v^t, respectively.

Generations do not overlap and each one reaches its end after the s time steps. Then, each agent u will have accumulated a total payoff of Π_u, corresponding to its fitness in a strong selection pressure process [18, 46]. Each agent lives for a generation and all of them reproduce synchronously by a localized *death–birth* process [47]: at the end of each generation, every node u will be replaced by a node u' in the set $\{u \cap N(u)\}$ which is composed by the neighbourhood of u ($N(u)$) and u itself (figure 1(c)). Accordingly, the network will stay the same for the whole simulation and changes are made only to the agents in each site. Node u' is chosen probabilistically according to the fitness ($\Pi_{u'}$) of nodes in N_2. Thus, on one hand, the nodes which accumulate more payoff are more likely to be chosen, on the other hand, even the most adapted agents can reproduce up to sites of distance one. Finally, some fluctuations might affect offspring. Specifically, there is a probability p_{mut} of a newborn having a bit flipped in their chromosome (figure 1(d)).

2.2. Game

We are interested in the evolution of cooperation in a population of agents facing a social dilemma. Strictly speaking, we want to check if cooperative heuristics are the most adapted in conditions wherein equilibria with pure strategies would be of full defection. We consider that at each interaction, agents play a round of a Prisoners' dilemma (PD) game with their neighbours. The PD is a 2×2 game in which only two actions are available to the players, either cooperate or defect. If two players cooperate, they both get a reward R, if one cooperates and the other defects, the cooperator earns S and the defector gets a payoff T (the

Table 1. Memory variables and genes. Genes determine agents actions, and they are responses to the variables stored in their memory. Here we show the variables considered and their corresponding gene for a previous round *l*. The description indicates how the information is stored and the role of each gene.

Variable	Gene	Description
	β_u^0	Constant response
$C_{v,u}^l$	$\mathbb{C}_u^l$	Direct reciprocity: 1 if v cooperated with u in round $t - l$, 0 otherwise
$R_{v,u}^l$	$\mathbb{R}_u^l$	Indirect reciprocity: fraction of times agent v cooperates in round $t - l$ with players other than u[a]
π_v^l	$\mathbb{P}_u^l$	Payoff obtained by agent v in round $t - l$
$D_{v,u}^l$	$\mathbb{D}_u^l$	Punishment: 0 if v cooperated with u in round $t - l$, 1 otherwise

[a] Typically indirect reciprocity is defined by individuals playing a one-shot game in a large well-mixed population [50]. Nonetheless, here $R_{u,v}^t$ does not consider the actions of v with respect to u, which should correspond to an analogous effect.

temptation to defect). Finally, if both defect, both of them obtain *P*. The PD occurs when the elements of the payoff matrix are such that $T > R > P > S$, which implies that a rational player should defect because, whatever your opponent does, the best (in terms of having larger payoff) is to defect. Henceforth, we consider that the values of each entry are a normalized version of Axelrod's tournament [48] values. Namely: $T = 1/\langle k \rangle$; $R = 0.6/\langle k \rangle$; $P = 0.2/\langle k \rangle$; $S = 0$. As mentioned before, for these values, the prediction is that under a replicator dynamics, the system ends up in full defection [18]. We also note that small changes in this parametrization would not affect our results, as they are robust for a broad range of the temptation (*T*) parameter (see appendix A).

2.3. Agents

Agents are hardwired, and their heuristics do not change in the course of one generation, which corresponds to their lifetime. Their heuristics are determined by their chromosomes and constitute a stochastic way to evaluate the variables stored in their memory and make a decision on whether to cooperate or not. Agents' memory stores the variables from previous interactions, and we assume their working memory is limited [49]. Hence, agents can only store a finite set of variables from the previous *m* rounds. Specifically, an agent *u* with the set of neighbours $N(u)$, has stored in its memory $\mathbf{M_u}$ variables for all $v \in N(u)$ and for all $l \in [1, m]$. Therefore, $\mathbf{M_u}$ is a matrix wherein each row contains the values stored for one neighbour, as shown in figure 1(a).

The heuristics evaluate each stored variable according to a specific gene in the chromosome. Therefore, the expression of each gene is a weight given to a variable containing some information influencing agents' decision making. The vector $\mathbf{E}_u$ carries the responses of an agent *u*, i.e., its expressed genes values. They are given by a two complement representation of the gene bits, therefore, they are integers from -128 to 127.[4] The vector E_u contains the responses to the variables plus a constant response (β^0). Table 1 shows the set of variables stored and their corresponding genes. They are a basic set of external characteristics that an elementary agent can observe. Thus, they constitute a reasonable set of variables to be taken into account by a somewhat minimal heuristic. Finally, whether or not an agent will cooperate is determined by the sigmoid function specified by

$$\rho_{u,v} = f(\mathbf{E}_u, \mathbf{X}_{u,v}) = \frac{1}{1 + e^{-\kappa(\mathbf{X}_{u,v} \cdot \mathbf{E}_u)}}, \tag{1}$$

where $\rho_{u,v}$ corresponds to the probability of agent *u* cooperating with agent *v*. $\mathbf{X}_{u,v} = (1) \oplus \mathbf{M}_{u,v}$ corresponds to a vector composed by the number 1 in the first position, followed by the memory variables as specified in table 1. κ (henceforth set to 0.05) provides the steepness of the curve, and it is chosen in such a way that if the dot product of both vectors is greater (*resp. lesser*) than 100, the probability should be approximately 1 (*resp. 0*), as illustrated by figure 1(b).

3. Results

We ran simulations for populations of 1024 agents connected on a lattice (LTT) with a von Neumann neighbourhood and on RRN with the same nodes' degree ($k = 4$) generated by the configuration model [51] We evolved the model for 5×10^5 generations, each with 100 rounds, for different values of the p_{mut} parameter. Results for memory between 0 and 5 are shown in figure 2. When $m = 0$, the agents' chromosome is composed of only the constant response (β^0) and strategies are reducible to mixed

---

[4] Hence, when a mutation occurs in a gene, from its expressed value it can be added/subtracted a random power of 2, or have its sign and value changed.

Figure 2. Cooperation thrives at low mutation values. Fraction of cooperative actions at the steady-state according to the mutation probability. Colours and shapes correspond for different memory (m) values. Averages plus 0.95 confidence interval of 100 realizations are presented for each mutation (p_{mut}) value.

Figure 3. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 1) and horizontal panels to different networks (LTT, RRN). Agents have memory $m = 1$ and 100 realizations are performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 generations.

strategies. In this case, when no mutation is available the system quickly goes to full defection (see figure D1), as expected, and mutation increases the possibility of adding cooperative strategies by drift. Conversely, when agents have access to memory, cooperation is predominant in the regime of low mutation. Furthermore, cooperation is larger and more resilient to higher mutation when agents have access to a bigger memory. With more memory, agents should be able to construct more complex heuristics which, here, seem to favour cooperation.

Figure 3 shows time evolution curves of individual realizations for $m = 1$ (results for other values of the memory are reported in appendix D). When $p_{mut} = 0$, the final cooperative fraction is highly dependent on the initial conditions, reaching a multitude of equilibria, some being fully cooperative and others showing a rather small level of cooperation, especially in the RRN network. Some executions have a distinguishably different fraction from the average, which can be a result of the finite system size. In the regime of small mutation rates, fluctuations increase significantly, however, for some small values of the mutation rate, all realizations converge to highly cooperative equilibria, as can be seen when $p_{mut} = 0.05$. Note, additionally, that as the probability of mutation increases, the fraction of cooperative actions decreases. For the limiting value $p_{mut} = 1$, every new player is born with a mutation, and the system evolves into a negligible average level of cooperation. Interestingly, this is a demonstration that a small noise can foster cooperation in the

Figure 4. Distribution of genes' expressed values. Densities of genes values for simulations on LTT and RRN graphs for $m = 1$. Top panels show distributions for $p_{\text{mut}} = 0.05$ and bottom panels for $p_{\text{mut}} = 1$. The vertical dashed line indicates separate regions wherein the marginal probability to cooperate would be smaller (negative gene values) and greater (positive gene value) than 0.5.

process of evolution. With more mutation, it gets harder for cooperative strategies to prevail and defection tends to increase, however, a sufficiently small mutation probability will guarantee that the system evolves to a cooperative equilibrium.

3.1. Heuristics and strategies

In this section, we focus on the composition of the populations in the different regimes. It is not straightforward to evaluate how genes and variables interact, hence, it is hard to determine if agents are going to cooperate or not in a specific situation. A first step is to investigate what are the gene values in cooperative and non-cooperative equilibria. Figure 4 show the distributions of genes for two mutation values: $p_{\text{mut}} = 0.05$ and $p_{\text{mut}} = 1$, wherein evolution leads to mostly cooperation and to mostly defection, respectively. Simulations in both LTT and RRN networks yielded similar distributions, indicating the presence of a common evolutionary pattern.

When the majority of the population cooperates ($p_{\text{mut}} = 0.05$), β^0, $\mathbb{C}^1$, and $\mathbb{R}^1$ have a clear right-modality with most of these values being higher than 0. Conversely, $\mathbb{D}^1$ is left-modal with a clear peak at extreme negative values, while $\mathbb{P}^1$ shows a softer trend towards negative values. This implies that when cooperation thrives, agents have a baseline cooperative response and tend to reciprocate cooperation both directly and indirectly. On the other hand, the agents punish defectors rigorously and have a mild negative response to other agents' payoff, probably as a means to punish defectors, as only defectors can attain the highest payoffs. Interestingly, the distributions of β^0 indicate that the emerging strategies are willing to cooperate even in a one-shot game (see figure E1) with an unknown player, albeit this is not the expected behaviour for $m = 0$. In the other extreme, for $p_{\text{mut}} = 1$, defection prevails, and genes values indicate the underpinnings of this trend. All distributions are right-skewed, with β^0 and $\mathbb{D}^1$ having a noticeable peak at the lowest possible values. Thus, when mutations are too frequent agents are much more likely to exploit and punish, leading defection to be the default strategy. Too much drift will make it impossible for cooperative heuristics to be selected, and they will vanish in the population.

These last results provide a picture of the genotype space. However, there is still the need to identify which strategies have emerged. When studying evolutionary games, it is always challenging to bridge the gap between the genotype and phenotype spaces [40]. In our model, the profile of agents' actions would correspond to observable phenotypes, yet it is not straightforward to specify a method for heuristics classification. An unsupervised procedure would fall into the problem of how to identify the groups encountered, i.e., how to determine to which known strategies they correspond. Therefore, here we adopted an approach that consisted of classifying agents by looking at what would be their responses to the most basic strategies: a pure defector and a pure cooperator. Namely, we looked at whether agents were likely to cooperate or defect with agents having a history corresponding to each of the two pure strategies. For instance, a full defector v would always have defected with u ($C_{v,u}^1 = 0$, $D_{v,u}^1 = 1$), with its other neighbours ($R_{v,u}^1 = 0$), and would have an expected payoff (π_v^1) corresponding to these actions.

Table 2. Classification of heuristics according to their responses to the two pure strategies: pure defector and pure cooperator. We consider that agents cooperate (*C*) or defect (*D*) if their probability to cooperate is greater than $(1 - \sigma)$ or smaller than σ, respectively.

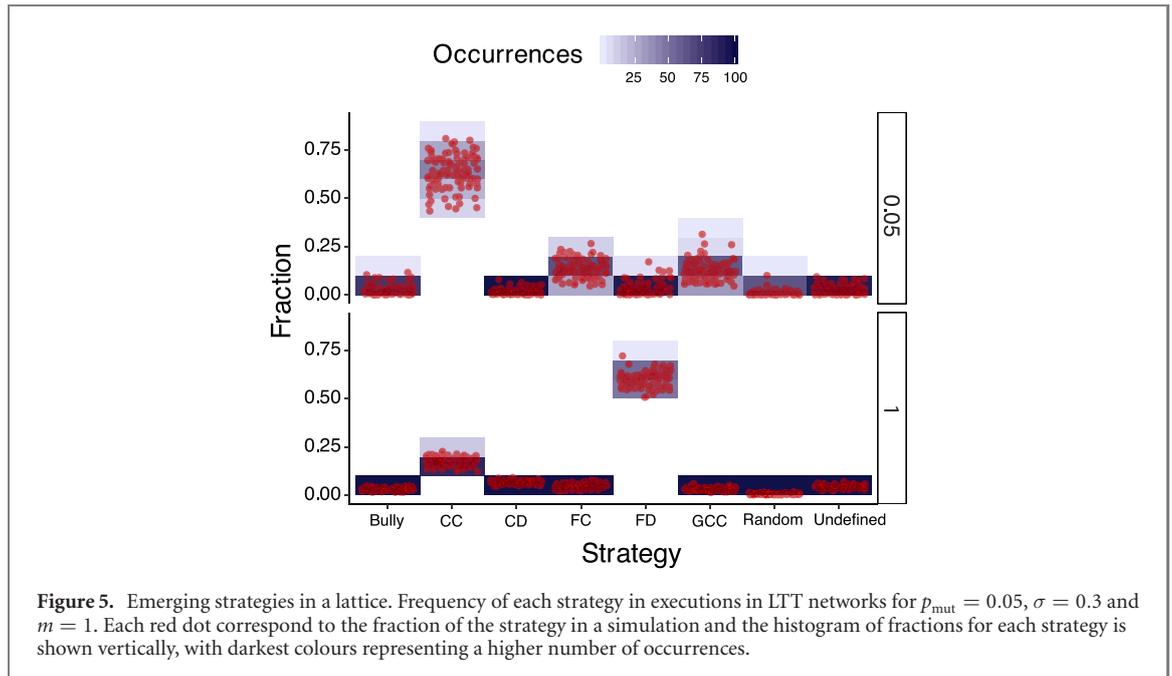|	Pure cooperator	Pure defector
FC	*C*	*C*
FD	*D*	*D*
CC	*C*	*D*
GCC	*C*	—
CD	—	*D*
Bully	*D*	*C*
Random	—	—

Figure 5. Emerging strategies in a lattice. Frequency of each strategy in executions in LTT networks for $p_{\mathrm{mut}} = 0.05$, $\sigma = 0.3$ and $m = 1$. Each red dot correspond to the fraction of the strategy in a simulation and the histogram of fractions for each strategy is shown vertically, with darkest colours representing a higher number of occurrences.

The proposed classification is shown in table 2 (see the details in appendix B). We considered strategies analogous to known ones, namely: *full Cooperator (FC)*, cooperates with both pure cooperators and pure defectors; *full Defector (FC)*, defects with both; *CC*, reciprocates cooperation and defects otherwise; *generous conditional cooperator (GCC)*, reciprocates cooperation and can cooperate randomly with defectors; *conditional defector (CD)*, cooperates randomly with cooperators and always defects with defectors; *bully*, defects with cooperators, but cooperates with defectors; *random*, behave randomly with both pure strategies. We labelled agents that could not be classified by this process as *undefined*.

In figure 5, we show the frequencies of each strategy from simulations of the heuristics selection dynamics on a lattice (a similar pattern is obtained for RRN networks, see figure B1). When the mutation is low ($p_{\mathrm{mut}} = 0.05$), most of the agents tend to be cooperators or CCs (mean fraction is 0.9 with a standard deviation of 0.07): CC constitutes most of the strategies, followed by a small fraction of GCC and FC players. In contrast, when mutation is high ($p_{\mathrm{mut}} = 1$), FD and CD constitute the majority (mean = 0.66, sd = 0.038) of agents. However, a minority of CC players can persist (mean = 0.17, sd = 0.022), which explains the existence of a small fraction of cooperative actions even in this regime.

4. Exploring kin discrimination: a first extension

It is known that cooperative behaviour can emerge and be sustained by factors that do not depend on players history of decisions. Namely, genetic relatedness or kinship plays a key role in the evolution of cooperation in nature [37, 44, 45, 52]. Kin selection is pervasive [4, 5], despite controversies over its role in particular phenomena [53–58]. Indeed, these disagreements indicate the need to investigate the role played by genetic relatedness in each specific scenario [56]. Therefore, to address this question, we take such mechanisms into account in the evolutionary dynamics of heuristics selection. Namely, we have extended the previous analysis and considered that agents could evaluate an additional variable that accounts whom
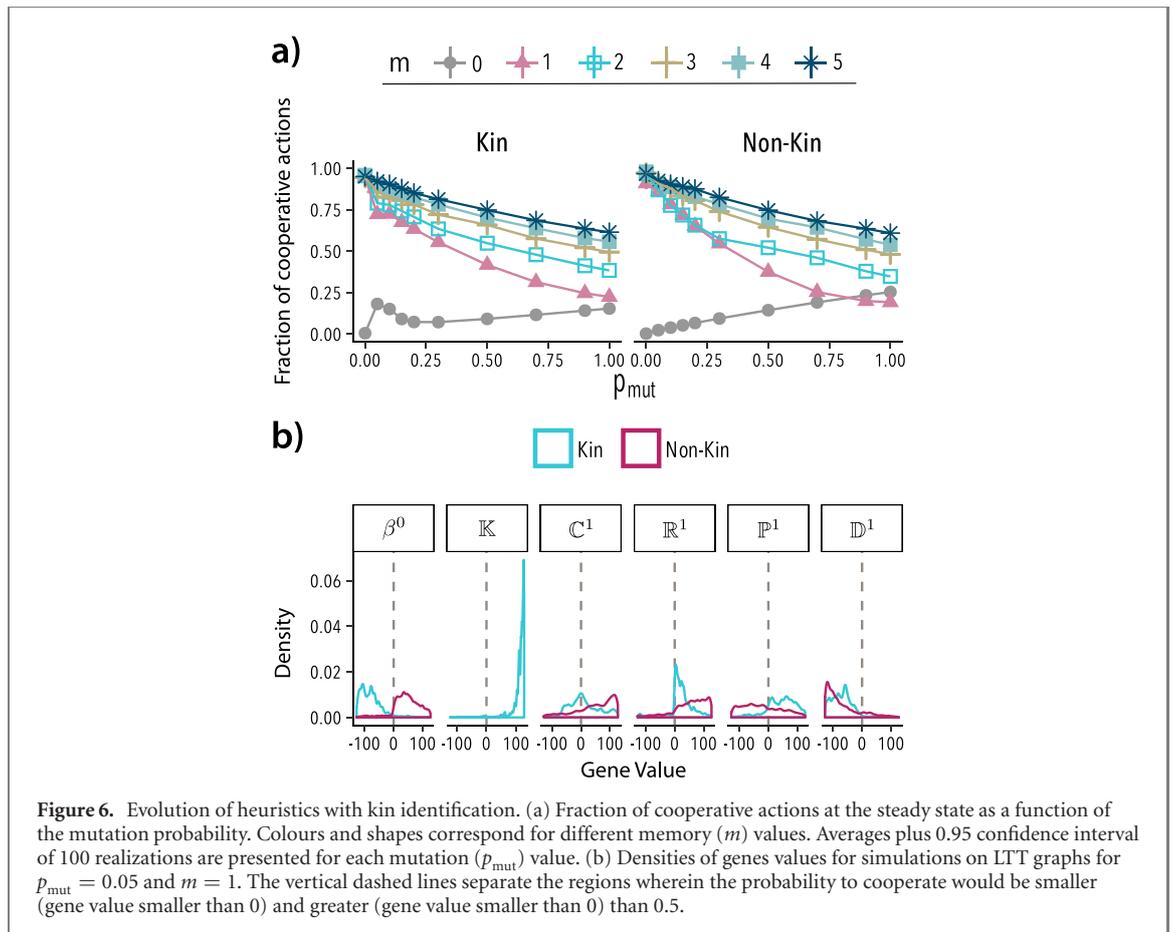
Figure 6. Evolution of heuristics with kin identification. (a) Fraction of cooperative actions at the steady state as a function of the mutation probability. Colours and shapes correspond for different memory (m) values. Averages plus 0.95 confidence interval of 100 realizations are presented for each mutation (p_{mut}) value. (b) Densities of genes values for simulations on LTT graphs for $p_{mut} = 0.05$ and $m = 1$. The vertical dashed lines separate the regions wherein the probability to cooperate would be smaller (gene value smaller than 0) and greater (gene value smaller than 0) than 0.5.

they are interacting with, specifically, genetic proximity, which is one possible mechanism ensuring interactions occur among related individuals [44].

   We added to the agents' chromosome a gene $\mathbb{K}$ to account for genetic relatedness with the interacting agent. Operationally, we consider that this kinship relation is given by the Jaccard index of pairs of agents' chromosomes. Note that we are not specifying a method for kin selection, but allowing the heuristics to take into consideration agents similarity when deciding to cooperate or not. Enabling, thus, an estimation of the relevance of genetic relatedness by evaluating the weight organically given to the heuristics' new gene.

   Results of simulations on a lattice are presented in figure 6. Figure 6(a) shows the fraction of cooperation at the steady-state both for our previous model (*non-kin*) and for the extended model (*kin*). The evolution leads to similar scenarios in both cases, indicating that the presence of the ($\mathbb{K}$) gene did not enhance nor undermine cooperation significantly, though there is one modest exception. For heuristics without memory ($m = 0$) and low mutation, there is a modest increase in the level of cooperation.

   Despite the negligible differences in outcomes, there is a substantial effect on agents' chromosomes. Figure 6(b) shows that including the possibility to weigh gene similarity changes the values of all other genes significantly. For $m = 1$, cooperation is strongly determined by the ($\mathbb{K}$) gene, and genes for direct reciprocity and constant response becomes negative or neutral. The latter implies that most agents will not cooperate in one-shot interactions with unrelated individuals, as shown in figure E2, demonstrating a significant difference from the agents without the $\mathbb{K}$ gene. There still is a mostly positive response for indirect reciprocity and a negative for punishment, while the weight given to participants payoff inverts. This result points to a compelling message: when heuristics can evaluate genetic relatedness, the ones that do that will have a higher reproduction, therefore resulting in more adapted heuristics. Nonetheless, information from past interactions is still required, with punishment and reciprocity playing a role.

5. Conclusions

Natural selection has shaped the evolution of all sort of life forms. Advantageous strategies endure while others dwindle in a never-ending process of adaptation. Accordingly, fundamental questions regarding the emergence of cooperative behaviour in social dilemmas have to be studied in the light of evolutionary mechanisms. Emerging behaviour is undoubtedly intrinsically dependant on the individuals under study,

e.g., humans commonly cooperate in large societies composed of unrelated individuals, while groups of animals are hardly greater than a few hundreds [59]. In particular, variance in humans is especially relevant, as behaviour is deeply affected by the specifics of the interactions and the culture of the individuals [23, 60]. Moreover, given that it is an emergent phenomenon, behaviour can be deeply affected by the complex topology of interactions [61]. In an attempt to provide a framework for such scenarios, here we explore a model that allows unravelling what could be the drivers of cooperation by a heuristics selection process.

By exploring heuristics that make use of agents behavioural information to stochastically determine their decisions in iterated PD games across generations, we have shown that, in a feasible environment, evolution will drive heuristics towards cooperation even when defection is expected for pure strategies. In these scenarios, reciprocity and punishment are the main ingredients of cooperators' decision-making, and most strategies will follow conditional cooperation. The fraction of cooperative decisions decreases with an increase in the mutation rate, nonetheless, for small mutation rates the system reaches a cooperative equilibrium. Without mutation, the configuration of the initial state is critical and the system can get trapped in equilibria of meagre cooperation. Increasing the memory of individuals also increases the fraction of cooperation, suggesting that heuristics with more resources are more cooperative. These aggregate results are indistinguishable from a version of the model wherein agents have, in addition to behavioural information, access to their similarity with others (which mimics genetic relatedness). For this latter scenario, the level of cooperation at the macroscopic level remains roughly the same. Important enough, however, at the level of individuals, chromosomes change significantly and cooperation is given through a kin identification process.

Therefore, when agents discriminate their kin, reciprocity loses much of its importance, which is especially insightful given the behaviour observed in nature. Kin selection is arguably the most important mechanism behind cooperation in non-human animals, while reciprocity is uncommon [44, 45]. Our result suggests that in order for reciprocity to be dominant, perfect kin discrimination cannot exist, which suggests that figuring out the interplay between both mechanisms is crucial for understanding human evolution. Moreover, agents evolved in each condition presented a different expected response in one-shot games with unrelated individuals: cooperation is likely without the kin discrimination gene, while the majority of agents will defect when they can discriminate their genetic similarity.

To round off, we note that heuristics will adapt according to the information that they have access to, and they can change significantly according to the variables available. Surprisingly, despite changes in methods, cooperation is more likely than exploitation, due to reciprocity [62, 63] or to kin selection [37]. This suggests that even if individuals have limited cognitive capacities (a small memory weighed by a rather inexpensive function), cooperative heuristics can have higher reproduction rates and be pervasive. However, extrapolations have to be made with caution. As it is often the case of works in evolutionary game theory, our model sidesteps important details from biology and cognitive sciences [64]. Besides, the presented results concern only regular graphs with a fixed degree and further explorations are necessary to assess how much our findings are constrained to the topologies studied here. Furthermore, future work should also explore the intersection between moral and material values and how it influences the heuristics [6], and how selection works in more complex scenarios, for instance, when higher cognition has higher associated costs [65]. In this line, our approach could be applied to understand how cultural characteristics [23, 35] drive cooperation in different directions by modelling proper environmental variables, and whether costly punishment could sustain large scale cooperation [66]. We plan to explore this and similar questions next.

Code availability

The code used in the simulations is available at: https://github.com/fmcardoso/heuristics_selection.

Acknowledgments

Figure A1. Average cooperation at the stationary state. The left panel shows results for $m = 0$ and the right panel for $m = 1$. 100 simulations were done for each p_m and T' combination. Colour coding corresponds to the average over all realizations and varies from blue (1) to yellow (0).

Figure A2. Distribution of genes' expressed values. Densities of genes values for simulations on a lattice for $m = 1$. Top panels show distributions for $p_{\mathrm{mut}} = 0.05$ and bottom panels for $p_{\mathrm{mut}} = 1$. The vertical dashed lines separate regions wherein the marginal probability to cooperate would be smaller (negative gene values) and greater (positive gene value) than 0.5.
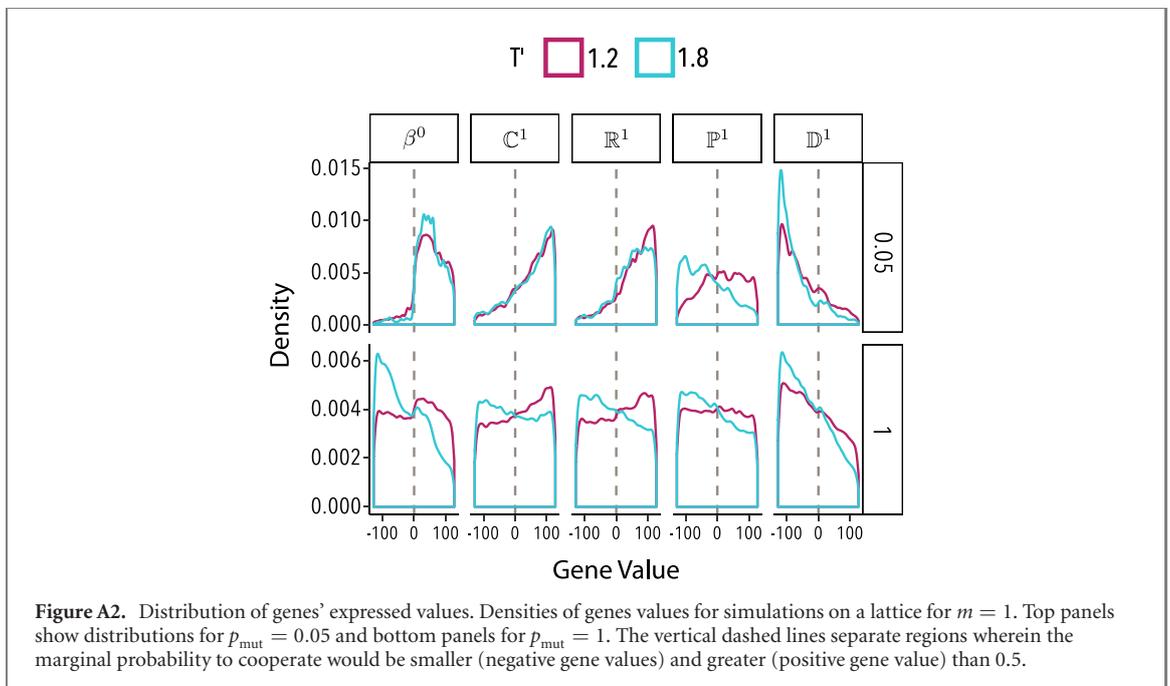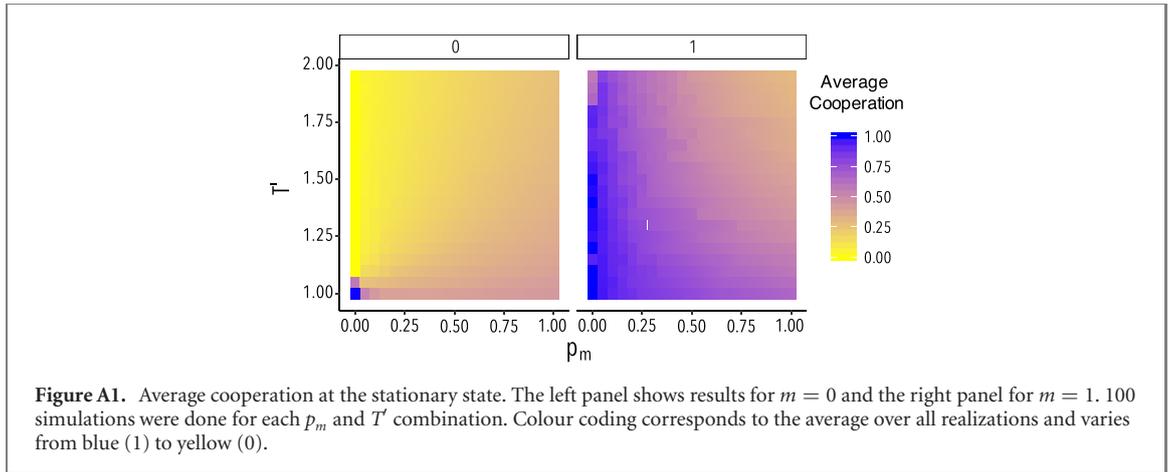
Appendix A. Other payoff values

To ensure that our results are robust with respect to differences in the payoff values, we ran simulations for different values of the temptation parameter T. To make our results comparable to previous work, we used the one-dimensional parametrization of payoffs used by Nowak and May [8]. In this version, $R = 1$, $P = \epsilon$, $S = 0$, and T varies from 1 to 2, with ϵ being a value close to zero. As we consider normalized versions, the payoff here is defined by $T = T'/\langle k \rangle$; $R = 1/\langle k \rangle$; $P = 0.01/\langle k \rangle$; $S = 0$, with T' varying from 1 to 2. Results for memory 0 and 1 are shown in figure A1. The results show that without memory, cooperation is only attainable when $T' = 1$ and low mutation. However, when agents have memory of their last interaction, cooperation endures even when the temptation to defect is around 2. As an illustration, the distributions of
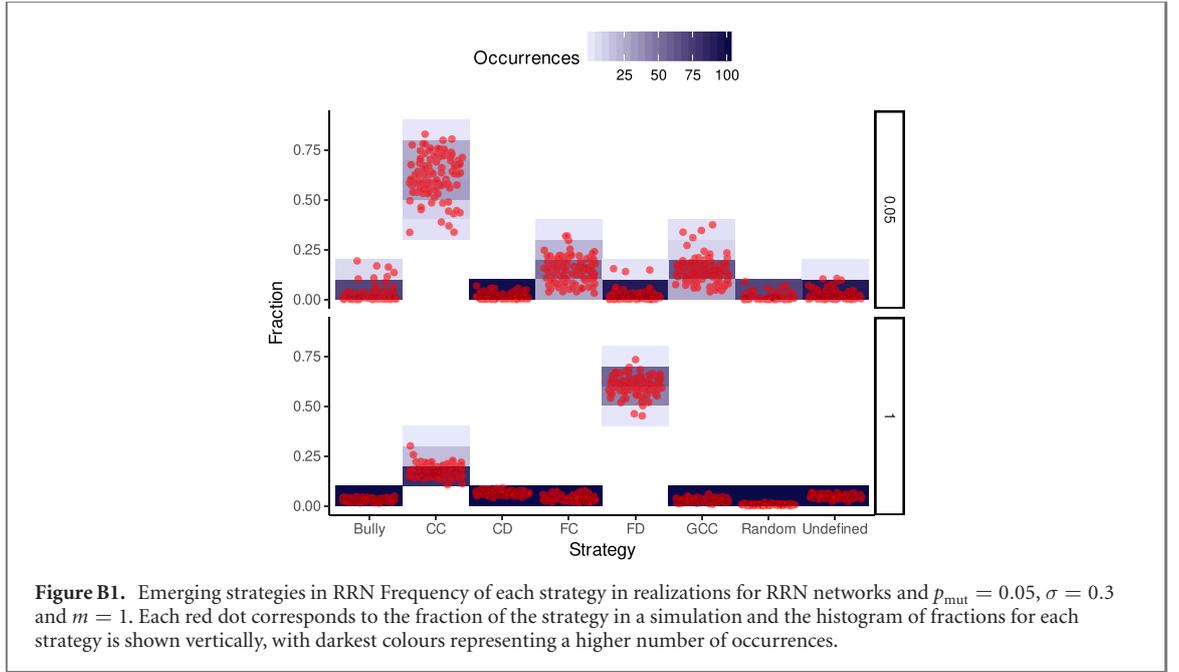
Figure B1. Emerging strategies in RRN Frequency of each strategy in realizations for RRN networks and $p_{\mathrm{mut}} = 0.05$, $\sigma = 0.3$ and $m = 1$. Each red dot corresponds to the fraction of the strategy in a simulation and the histogram of fractions for each strategy is shown vertically, with darkest colours representing a higher number of occurrences.

Table B1. Pure strategies memory. Past history of the two basic strategies according to what would have been played by them for $m = 1$.

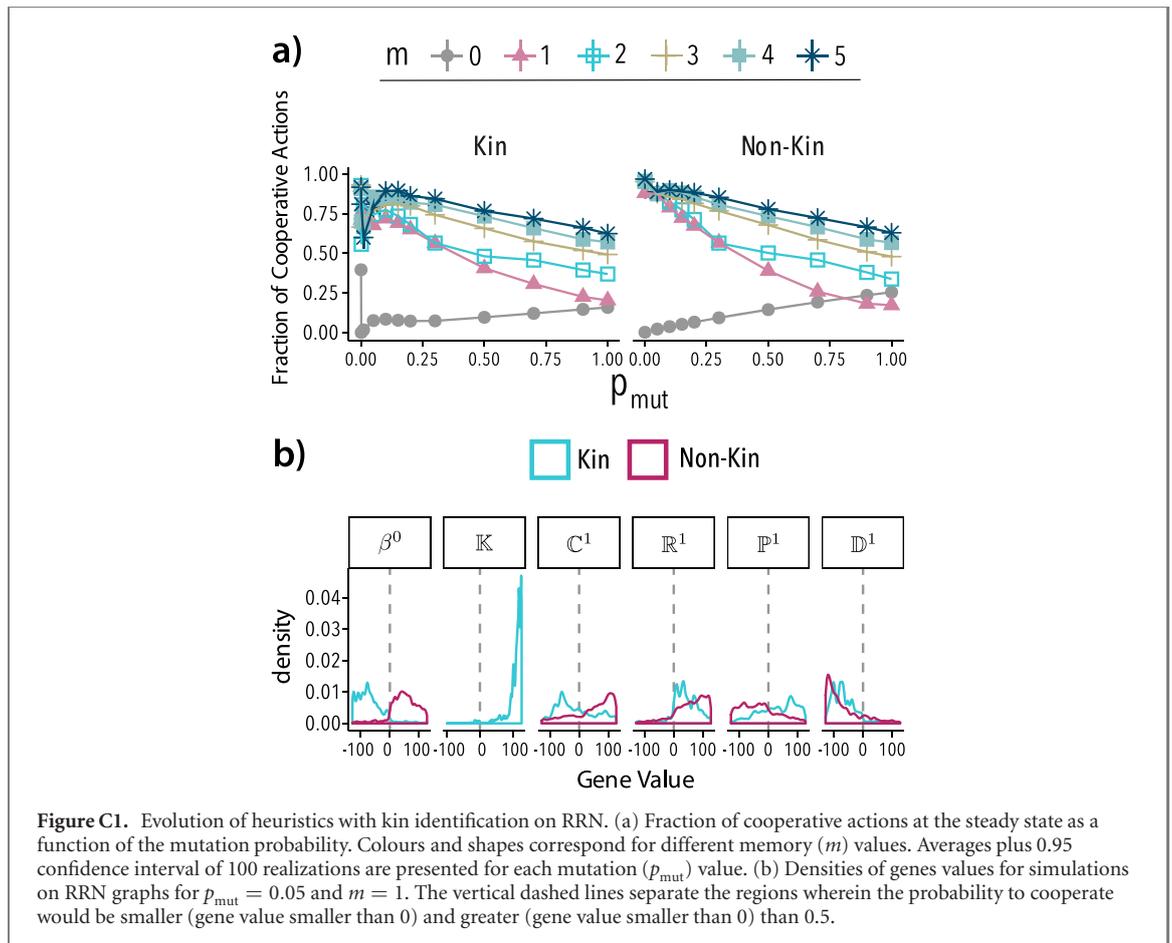|	Pure cooperator	Pure defector
$C^1_{v,u}$	1	0
$R^1_{v,u}$	1	0
π^1_v	$\langle \pi_C \rangle$	$\langle \pi_D \rangle$
$D^1_{v,u}$	0	1

Table B2. Classification of heuristics according to their responses to the two pure strategies: pure defector and pure cooperator. ρ corresponds to the resulting probability when playing with the corresponding pure strategy.

	Pure cooperator	Pure defector
FC	$\rho > (1 - \sigma)$	$\rho > (1 - \sigma)$
FD	$\rho < \sigma$	$\rho < \sigma$
CC	$\rho > (1 - \sigma)$	$\rho < \sigma$
GCC	$\rho > (1 - \sigma)$	$\sigma \leqslant \rho \leqslant (1 - \sigma)$
CD	$\sigma \leqslant \rho \leqslant (1 - \sigma)$	$\rho < \sigma$
Bully	$\rho < \sigma$	$\rho > (1 - \sigma)$
Random	$\sigma \leqslant \rho \leqslant (1 - \sigma)$	$\sigma \leqslant \rho \leqslant (1 - \sigma)$

gene values for T' values of 1.2 and 1.8 are shown in figure A2. They follow a close pattern to the ones shown in figure 4 in the main text, indicating the equivalence of both results. This shows that our results are robust across a broad range of parameter values.

Appendix B. Heuristics classification

Heuristics are classified according to two basic strategies: *pure cooperation* and *pure defection*. These two strategies always cooperate and always defect, respectively. Table B1 illustrates the variables contained in the memory of agent u with respect to a player v, corresponding to the two pure strategies for $m = 1$. All the values are given straightforwardly, expect for π^1_v. Payoffs values are more complicated, as they depend on the players with whom they are playing with, which we cannot define *a priori*. We decided to use the average payoff of individuals which cooperated and defected with all their neighbours for the pure cooperator and pure defector, respectively. Therefore, $\langle \pi_C \rangle = \langle \pi^t_i \rangle \forall\, i \in R_1, \forall\, t \in [1, 100]$ and $\langle \pi_D \rangle = \langle \pi^t_i \rangle \forall\, i \in R_0$, $\forall\, t \in [1, 100]$, wherein R_1 (resp. R_0) corresponds to the set of agents which cooperated with all (resp. none) of their neighbours in the last time step.

Figure C1. Evolution of heuristics with kin identification on RRN. (a) Fraction of cooperative actions at the steady state as a function of the mutation probability. Colours and shapes correspond for different memory (m) values. Averages plus 0.95 confidence interval of 100 realizations are presented for each mutation (p_{mut}) value. (b) Densities of genes values for simulations on RRN graphs for $p_{\text{mut}} = 0.05$ and $m = 1$. The vertical dashed lines separate the regions wherein the probability to cooperate would be smaller (gene value smaller than 0) and greater (gene value smaller than 0) than 0.5.

The activation function (equation (1) in the main text) of an agent results in the probability to cooperate with the pure cooperator (ρ_{C}) and the pure defector (ρ_{D}). We then, use the threshold σ to divide the plane ($\rho_{\text{C}}, \rho_{\text{D}}$). Namely, we designate as cooperation when $\rho > (1 - \sigma)$, defection as $\rho < \sigma$, and random when $\sigma \leqslant \rho \leqslant (1 - \sigma)$. This process results in the set of strategies given in table 2 of the main text. Therefore, a precise version of this table would correspond to the one shown in table B2.

Appendix C. Extended model in RRN

In this section, we present the results of the extended model with the Kinship parameter ran on RRN graphs (figure C1) At variance with the model in a lattice, when there is no mutation, the fraction of cooperative actions can be different from zero, as it is also shown in the time evolution figures. This demonstrates how important the kin identification mechanism can be in an adequate environment. With mutation, the macroscopic results are equivalent to the results in a lattice and in an RRN without the extension. Furthermore, when agents have had access to memory the results are equivalent to the ones obtained in a lattice, including the distribution of gene values for $m = 1$.

Appendix D. Time evolution

Time evolution curves of the main model executions for memory values of 0,2,3,4, and 5 are shown in figures D1–D5, respectively. Figures D6–D11 show the time evolution curves of realizations of the model with the addition of the kin identification gene for memory values of 0, 1, 2, 3, 4, and 5, respectively.

Appendix E. One-shot responses

Distributions of the probability of cooperating with an unknown agent in a one-shot game, are shown in figures E1 and E2 for the non-kin and kin models, respectively.

Figure D1. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 0$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D2. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 2$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D3. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 3$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D4. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 4$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.
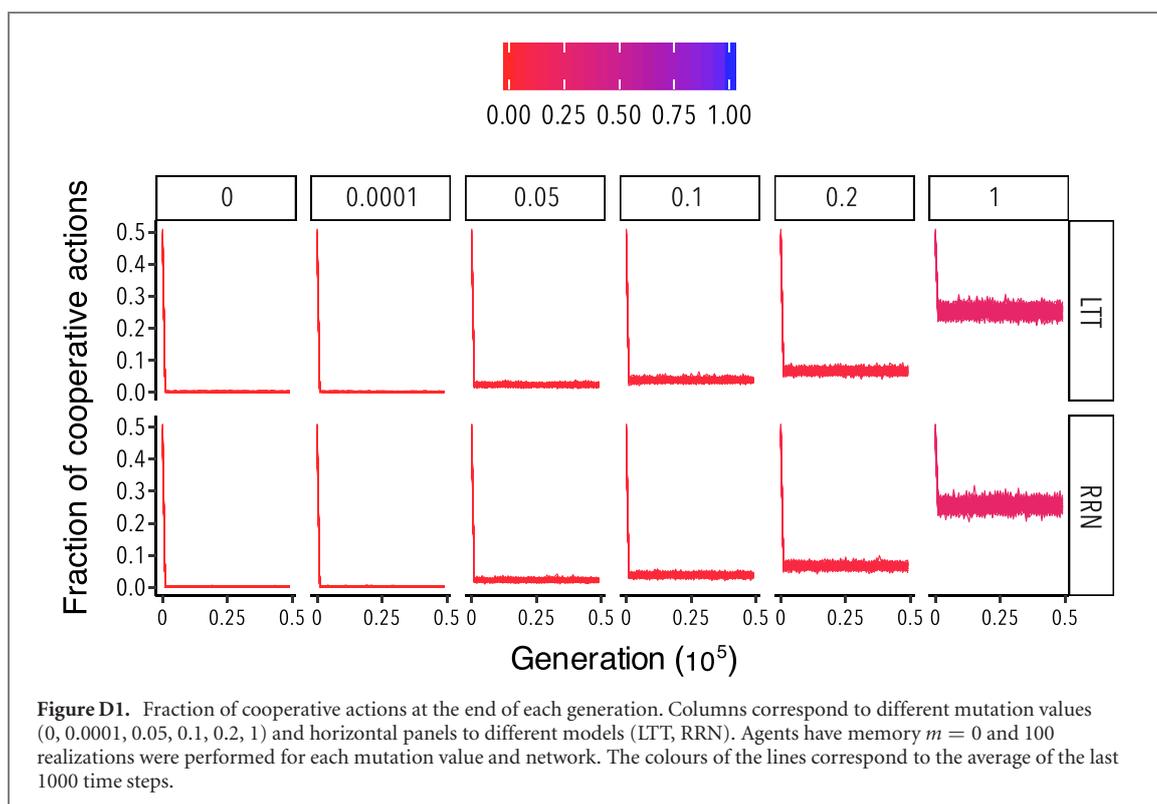
Figure D5. Fraction of cooperative actions at the end of each generation. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 5$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D6. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m =$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.
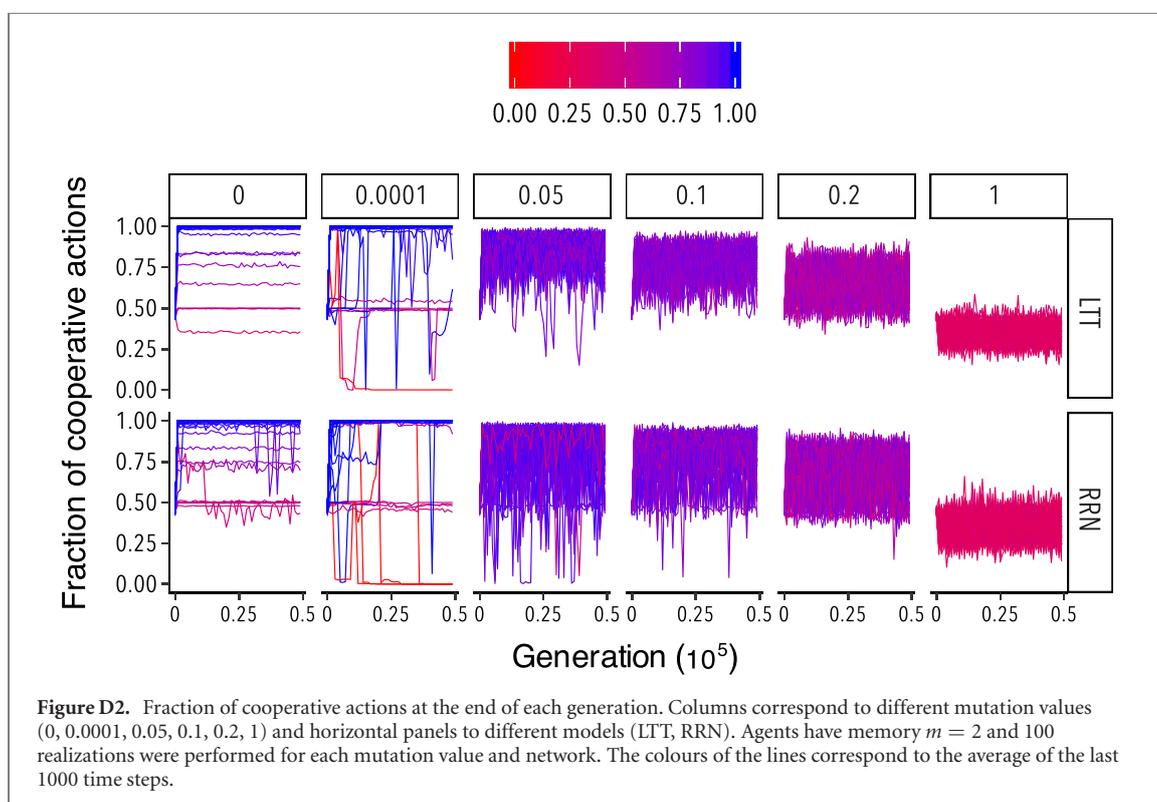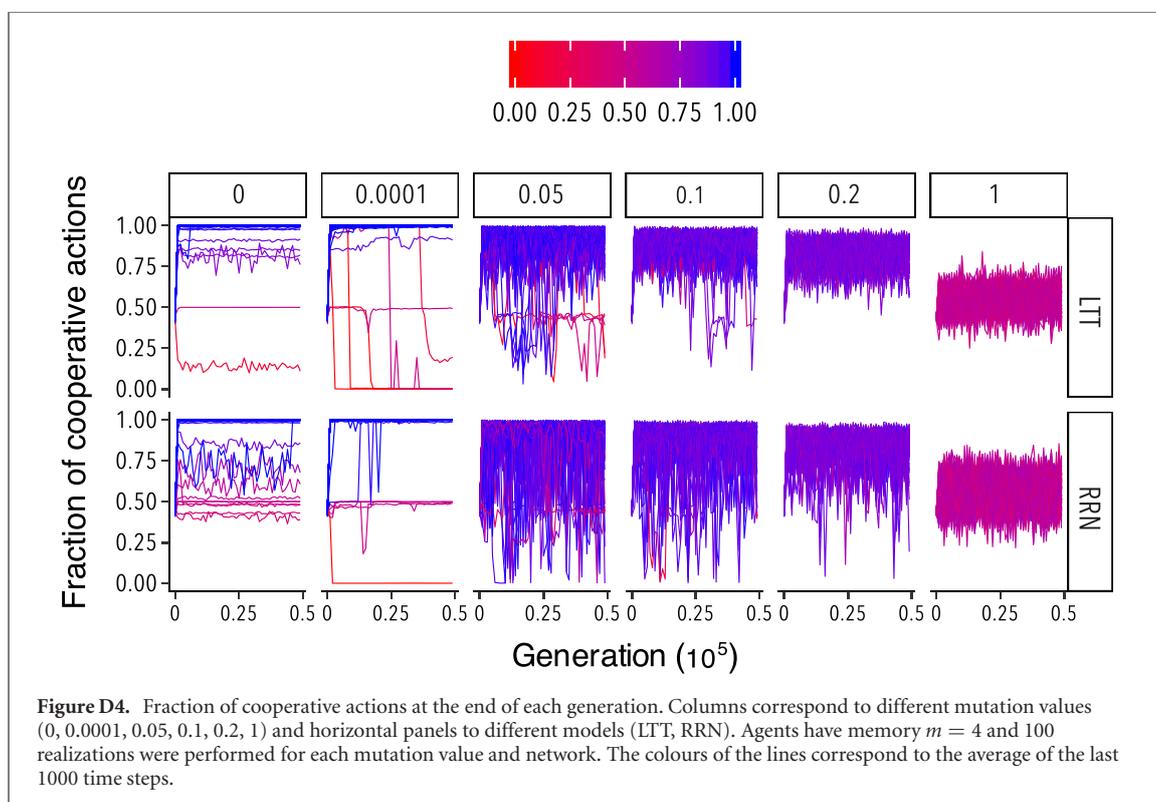
Figure D7. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 1$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.
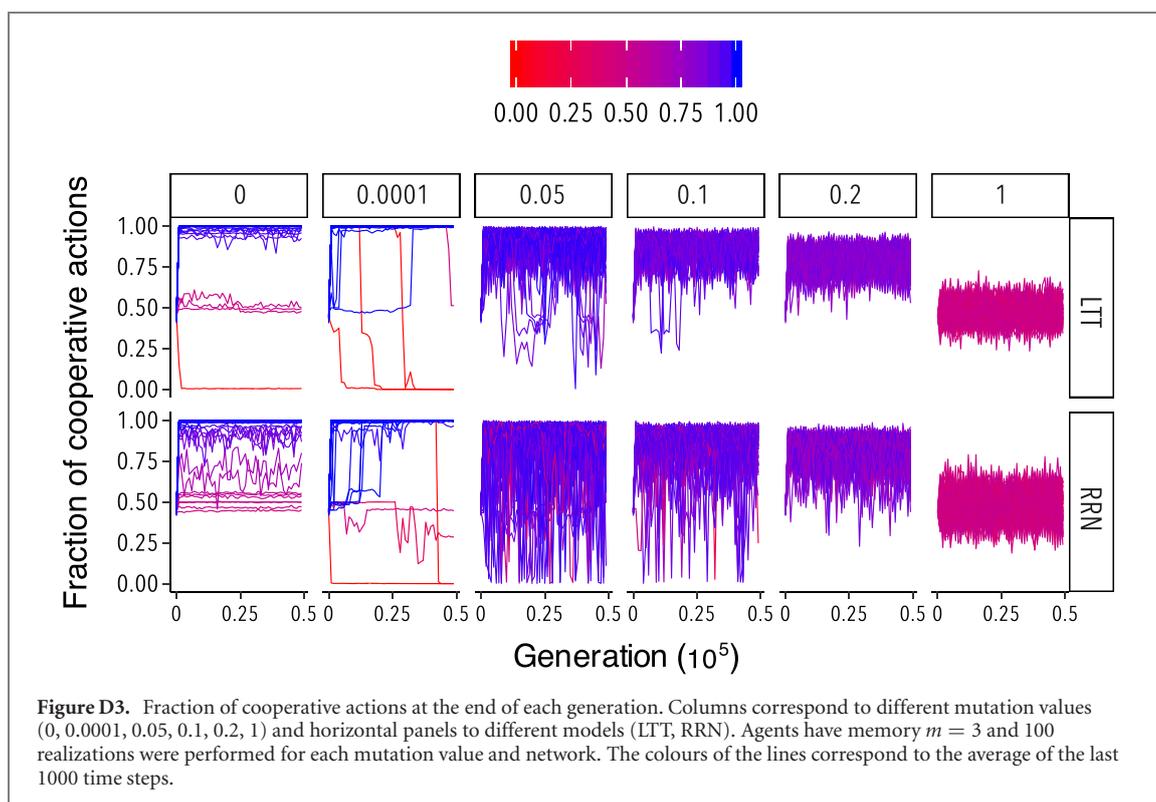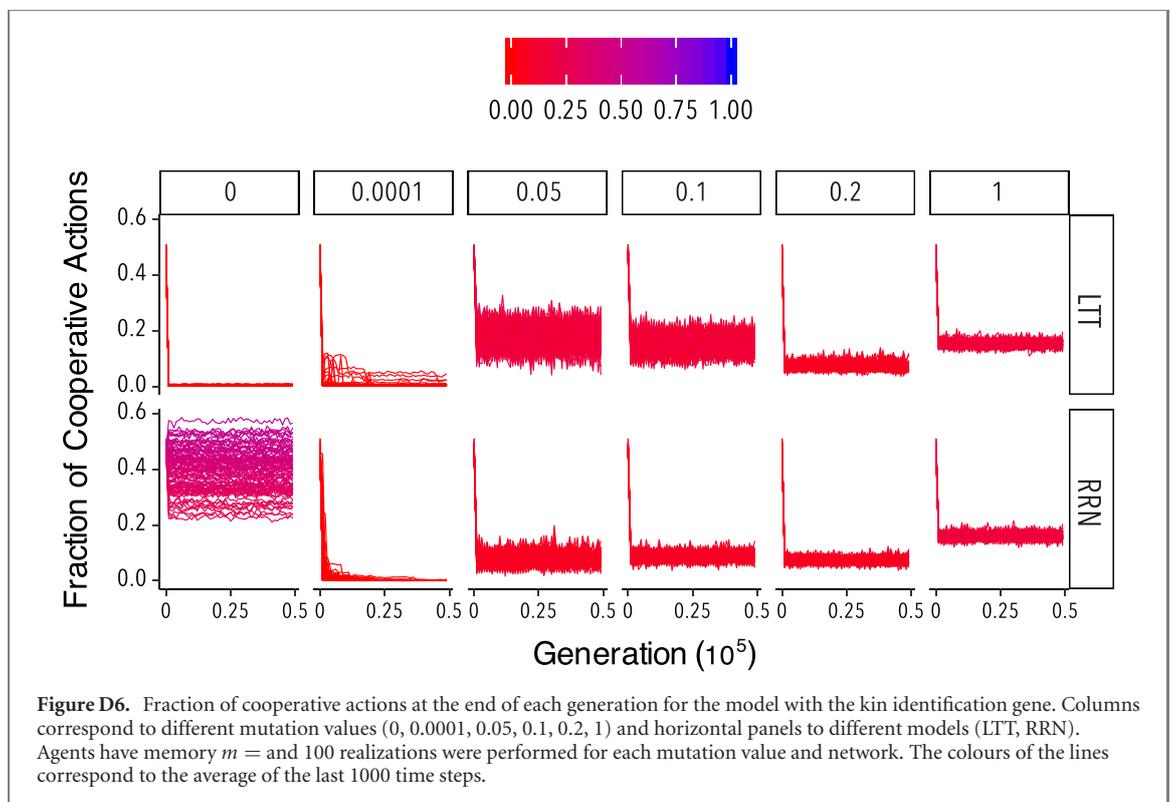

Figure D8. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 2$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D9. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 3$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.
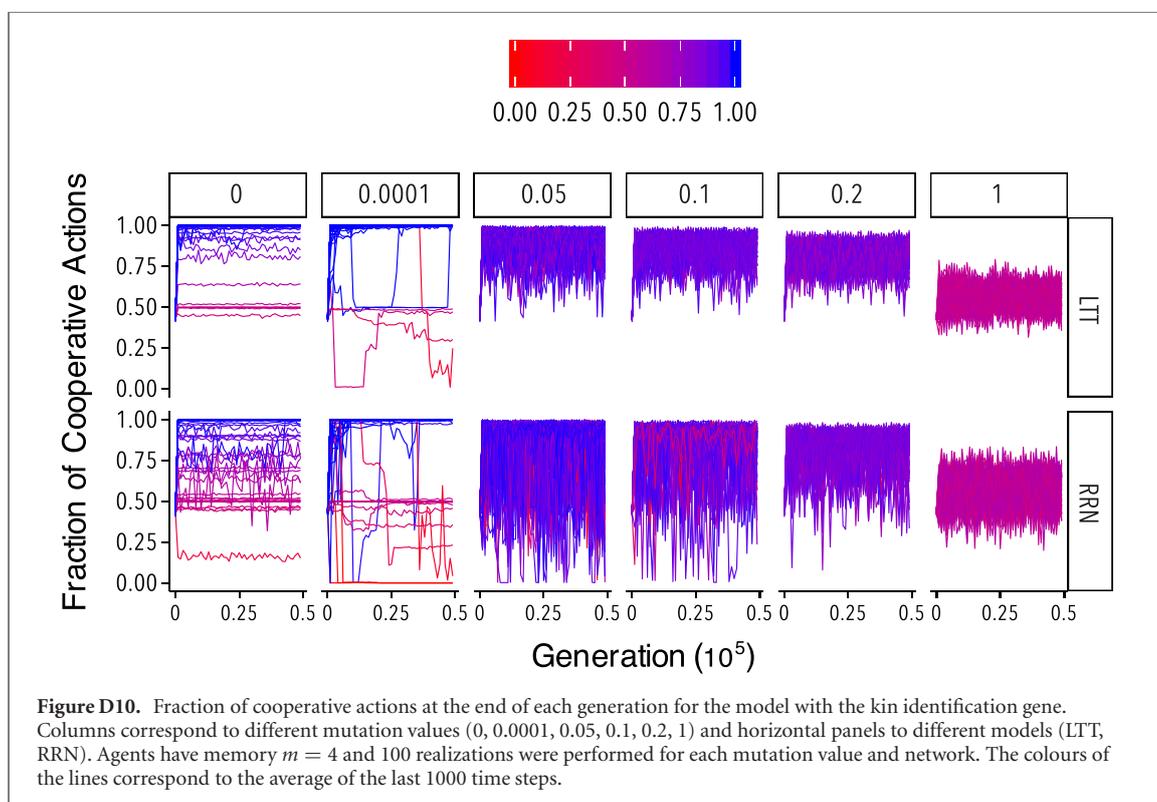

Figure D10. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 4$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.

Figure D11. Fraction of cooperative actions at the end of each generation for the model with the kin identification gene. Columns correspond to different mutation values (0, 0.0001, 0.05, 0.1, 0.2, 1) and horizontal panels to different models (LTT, RRN). Agents have memory $m = 5$ and 100 realizations were performed for each mutation value and network. The colours of the lines correspond to the average of the last 1000 time steps.
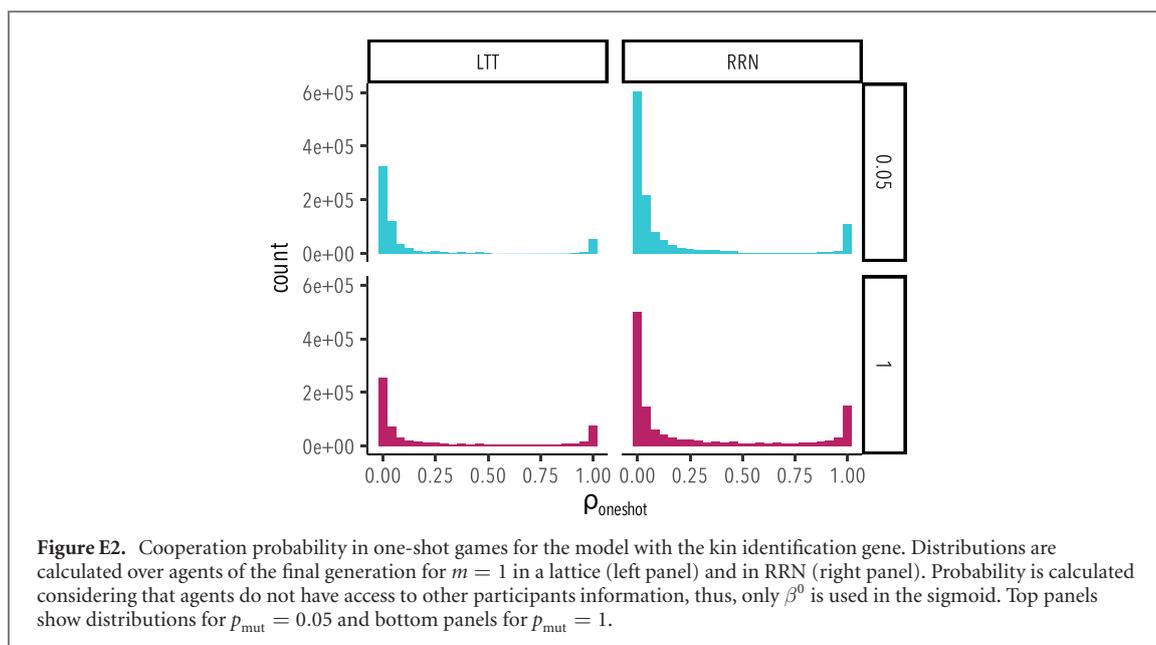

Figure E1. Cooperation probability in one-shot games. Distributions are calculated over agents of the final generation for $m = 1$ in a lattice (left panel) and in RRN (right panel). Probability is calculated considering that agents do not have access to other participants information, thus, only β^0 is used in the sigmoid. Top panels show distributions for $p_{\text{mut}} = 0.05$ and bottom panels for $p_{\text{mut}} = 1$.

Figure E2. Cooperation probability in one-shot games for the model with the kin identification gene. Distributions are calculated over agents of the final generation for $m = 1$ in a lattice (left panel) and in RRN (right panel). Probability is calculated considering that agents do not have access to other participants information, thus, only β^0 is used in the sigmoid. Top panels show distributions for $p_{mut} = 0.05$ and bottom panels for $p_{mut} = 1$.

ORCID iDs

Felipe Maciel Cardoso ⓘ https://orcid.org/0002-2789-9250
Carlos Gracia-Lázaro ⓘ https://orcid.org/0000-0002-9769-8796
Yamir Moreno ⓘ https://orcid.org/0000-0002-0895-1893

References

[1] von Neumann J, Morgenstern O, Kuhn H W and Rubinstein A 1944 *Theory of Games and Economic Behavior* (Princeton, NJ: Princeton University Press) p 776
[2] Myerson R B 1997 *Game Theory: Analysis of Conflict* 1st edn (Cambridge, MA: Harvard University Press)
[3] Rapoport A and Chammah A M 1965 *Prisoner's Dilemma: A Study in Conflict and Cooperation* (Ann Arbor, MI: University of Michigan Press)
[4] Dugatkin L A 1997 *Cooperation Among Animals: An Evolutionary Perspective* (Oxford: Oxford University Press)
[5] Bourke A F G 2011 *Principles of Social Evolution* (Oxford: Oxford University Press)
[6] Bowles S and Gintis H 2011 *A Cooperative Species: Human Reciprocity and its Evolution* (Princeton, NJ: Princeton University Press) https://press.princeton.edu/titles/9474.html
[7] Perc M, Jordan J J, Rand D G, Wang Z, Boccaletti S and Szolnoki A 2017 Statistical physics of human cooperation *Phys. Rep.* **687** 1–51
[8] Nowak M A and May R M 1992 Evolutionary games and spatial chaos *Nature* **359** 826–9
[9] Nowak M A and Sigmund K 1992 Tit for tat in heterogeneous populations *Nature* **355** 250–3
[10] Nowak M A 2006 Five rules for the evolution of cooperation *Science* **314** 1560–3
[11] Axelrod R 1987 The evolution of strategies in the iterated Prisoner's dilemma *Genetic Algorithms and Simulated Annealing* ed L Davis (Los Altos, CA: Morgan Kaufman) pp 32–41
[12] Lindgren K 1991 Evolutionary phenomena in simple dynamics *Artificial Life II* vol 10 pp 295–312 https://research.chalmers.se/en/publication/140676
[13] Szabó G, Vukov J and Szolnoki A 2005 Phase diagrams for an evolutionary Prisoner's dilemma game on two-dimensional lattices *Phys. Rev.* E **72** 047107
[14] Santos F C and Pacheco J M 2009 A new route to the evolution of cooperation *J. Evol. Biol.* **19** 726–33
[15] Vukov J, Szabó G and Szolnoki A 2006 Cooperation in the noisy case: Prisoner's dilemma game on two types of regular random graphs *Phys. Rev.* E **73** 067103
[16] Perc M and Szolnoki A 2008 Social diversity and promotion of cooperation in the spatial Prisoner's dilemma game *Phys. Rev.* E **77** 011904
[17] Perc M and Wang Z 2010 Heterogeneous aspirations promote cooperation in the Prisoner's dilemma game *Plos One* **5** e15117
[18] Roca C P, Cuesta J A and Sánchez A 2009 Evolutionary game theory: temporal and spatial effects beyond replicator dynamics *Phys. Life Rev.* **6** 208–49
[19] Colman A M, Browning L and Pulford B D 2012 Spontaneous similarity discrimination in the evolution of cooperation *J. Theor. Biol.* **299** 162–71
[20] Sigmund K 2010 *The Calculus of Selfishness* (Princeton, NJ: Princeton University Press)
[21] Javarone M A 2016 Statistical physics of the spatial Prisoner's dilemma with memory-aware agents *Eur. Phys. J.* B **89** 42
[22] Aleta A, Meloni S, Perc M and Moreno Y 2016 From degree-correlated to payoff-correlated activity for an optimal resolution of social dilemmas *Phys. Rev.* E **94** 1–8
[23] Henrich J *et al* 2005 Economic man in cross-cultural perspective: behavioral experiments in 15 small-scale societies *Behav. Brain Sci.* **28** 795–815
[24] Foster K and Ratnieks F 2005 A new eusocial vertebrate? *Trends Ecol. Evol.* **20** 363–4

[25] Traulsen A, Hauert C, De Silva H, Nowak M A and Sigmund K 2009 Exploration dynamics in evolutionary games *Proc. Natl Acad. Sci.* **106** 709–12

[26] Su Q, Li A, Zhou L and Wang L 2016 Interactive diversity promotes the evolution of cooperation in structured populations *New J. Phys.* **18** 103007

[27] Tversky A and Kahneman D 1974 Judgment under uncertainty: heuristics and biases *Science* **185** 1124–31

[28] Gintis H 2011 Gene–culture coevolution and the nature of human sociality *Phil. Trans. R. Soc.* B **366** 878–88

[29] Aumann R J 2019 A synthesis of behavioural and mainstream economics *Nat. Hum. Behav.* **3** 666–70

[30] Thaler R H 2016 Behavioral economics: past, present and future *SSRN Electronic Journal* **20** 9

[31] Knafo A *et al* 2008 Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS$_3$ promoter region and correlation between RS$_3$ length and hippocampal mRNA *Genes Brain Behav.* **7** 266–75

[32] Bateson M, Nettle D and Roberts G 2006 Cues of being watched enhance cooperation in a real-world setting *Biol. Lett.* **2** 412–4

[33] Rand D G, Greene J D and Nowak M A 2012 Spontaneous giving and calculated greed *Nature* **489** 427–30

[34] Rand D G, Peysakhovich A, Kraft-Todd G T, Newman G E, Wurzbacher O, Nowak M A and Greene J D 2014 Social heuristics shape intuitive cooperation *Nat. Commun.* **5** 3677

[35] Boyd R 2018 *A Different Kind of Animal: How Culture Transformed Our Species* (Princeton, NJ: Princeton University Press)

[36] Richerson P *et al* 2016 Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence *Behav. Brain Sci.* **39** e30

[37] Hamilton W D 1964 The genetical evolution of social behaviour. I *J. Theor. Biol.* **7** 1–16

[38] Eiben A E and Smith J 2003 *Introduction to Evolutionary Computing* (Berlin: Springer)

[39] Richerson P J and Boyd R 2005 *Not by Genes Alone: How Culture Transformed Human Evolution* (Chicago, IL: University of Chicago Press)

[40] Nowak M A 2004 Evolutionary dynamics of biological games *Science* **303** 793–9

[41] Miller J H 1996 The coevolution of automata in the repeated Prisoner's dilemma *J. Econ. Behav. Organ.* **29** 87–112

[42] van Veelen M, García J, Rand D G and Nowak M A 2012 Direct reciprocity in structured populations *Proc. Natl Acad. Sci.* **109** 9929–34

[43] Binmore K G and Samuelson L 1992 Evolutionary stability in repeated games played by finite automata *J. Econ. Theor.* **57** 278–305

[44] West S A, Griffin A S and Gardner A 2007 Evolutionary explanations for cooperation *Curr. Biol.* **17** R661–72

[45] Clutton-Brock T 2009 Cooperation between non-kin in animal societies *Nature* **462** 51–7

[46] Traulsen A, Pacheco J M and Nowak M A 2007 Pairwise comparison and selection temperature in evolutionary game dynamics *J. Theor. Biol.* **246** 522–9

[47] Lindgren K and Nordahl M G 1994 Evolutionary dynamics of spatial games *Physica* D **75** 292–309

[48] Axelrod R and Hamilton W 1981 The evolution of cooperation *Science* **211** 1390–6

[49] Milinski M and Wedekind C 1998 Working memory constrains human cooperation in the Prisoner's dilemma *Proc. Natl Acad. Sci.* **95** 13755–8

[50] Nowak M A and Sigmund K 2005 Evolution of indirect reciprocity *Nature* **437** 1291–8

[51] van der Hofstad R 2017 *Random Graphs and Complex Networks* (*Cambridge Series in Statistical and Probabilistic Mathematics* vol 1) (Cambridge: Cambridge University Press)

[52] Queller D C 1992 A general model for kin selection *Evolution* **46** 376–80

[53] Nowak M A, Tarnita C E and Wilson E O 2010 The evolution of eusociality *Nature* **466** 1057–62

[54] Nowak M A, McAvoy A, Allen B and Wilson E O 2017 The general form of Hamilton's rule makes no predictions and cannot be tested empirically *Proc. Natl Acad. Sci.* **114** 5665–70

[55] Abbot P *et al* 2011 Inclusive fitness theory and eusociality *Nature* **471** E1–4

[56] Birch J 2017 The inclusive fitness controversy: finding a way forward *R. Soc. Open Sci.* **4** 170335

[57] Birch J 2019 Are kin and group selection rivals or friends? *Curr. Biol.* **29** R433–8

[58] van Veelen M 2018 Can Hamilton's rule be violated? *eLife* **7** e41901

[59] Moffett M W 2013 Human identity and the evolution of societies *Hum. Nat.* **24** 219–67

[60] Camerer C F 2003 *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton, NJ: Princeton University Press) https://press.princeton.edu/books/hardcover/9780691090399/behavioral-game-theory

[61] Szabó G and Fáth G 2007 Evolutionary games on graphs *Phys. Rep.* **446** 97–216

[62] Trivers R L 1971 The evolution of reciprocal altruism *Q. Rev. Biol.* **46** 35–57

[63] Alexander R D 1987 *The Biology of Moral Systems* (New York: Routledge) https://taylorfrancis.com/books/9781351329309

[64] van Cleve J 2020 Building a synthetic basis for kin selection and evolutionary game theory using population genetics *Theor. Popul. Biol.* **133** 65–70

[65] Seoane L F and Solé R V 2018 Information theory, predictability and the emergence of complex life *R. Soc. Open Sci.* **5** 172221

[66] Fehr E and Gächter S 2002 Altruistic punishment in humans *Nature* **415** 137–40