ECAI 2024
U. Endriss et al. (Eds.)

doi:10.3233/FAIA240902

Counterfactual Thinking in Stochastic Dynamics of Cooperation

António M. Fernandes[a,*], **Francisco C. Santos**[a] **and Ana Paiva**[a,1]

[a]INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Abstract. Understanding the conundrum of human cooperation has been declared one of the century's grand challenges. Traditionally, the evolution of cooperative action in nature is analyzed through the lens of Evolutionary Game Theory, specifically, using the social learning framework, a model for Darwinian competition. However, more complex individuals may resort to more sophisticated learning rules such as Counterfactual Thinking (CT). Given these individuals' cognitive empowerment, the question of how the presence of counterfactuals influences the evolution of cooperation in a hybrid population of these complex agents and social learners. Here we explore how cooperation emerges from the interplay of different strategy revision paradigms by analyzing large-scale Markov processes. We find that increasing the prevalence of CT individuals can promote cooperation, but such an increase is non-monotonous. Moreover, whereas counterfactual reasoning generally fosters cooperation, it fails to promote such behaviour among counterfactuals. Lastly, we find that increasing the population's heterogeneity level enhances cooperation among social learners, but again not among counterfactuals. This indicates that, under certain circumstances, the presence of more sophisticated agents may help promote cooperation in hybrid populations. The proposed study may come as a starting point for a more profound understanding of agents' counterfactual rationality impact on hybrid populations.

1 Introduction

Cracking the enigma of collaborative behaviour in the natural world has been deemed one of the grand challenges of the century. It has been intriguing the scientific community in various areas such as sociology, biology, mathematics, physics and others [1, 18, 23, 24, 29, 30, 31, 34, 39]. The emergence and sustainability of cooperation at different levels of organization have been widely studied and modelled through the lens of evolutionary game theory, where special attention has been given to social dilemmas, either on its dyadic processes [7, 35, 38] or collective action [2, 12, 14, 15, 33, 41, 42].

To explain the emergence of cooperation, researchers have proposed many dynamic mechanisms, such as preferential selection [8, 44], punishment and reward [11, 29], memory effects [20, 50], and reputation [21, 25], among many other. Most studies assume that the decision process occurs through imitation and that decision-makers are purely rational [22]. Therefore, the population's dynamics have its framework based on the typical social learning (SL) definition from Evolutionary Game Theory (EGT), which assembles the base model for Darwinian competition [22, 40]. However, not only the pursuit of the highest payoffs in evolutionary social dilemmas is risky, but also individual's reasoning may not always be driven by the basic selfish agenda of conventional fitness maximization [43]. In fact, in many real-world scenarios, individuals may adopt different reasoning mechanisms, not being entirely rational for instance, as in the case of conformity-driven individuals who imitate the most abundant strategy in the population [17, 43, 44], or not complying with the imitation process, as in the case of counterfactual reasoning, where agents assemble their reasoning introspectively [5, 19]. The latter mechanism is of foremost interest, as it regards a very sophisticated ability, uniquely attributed to humans, that could have a major impact on future hybrid societies of agents and humans [27, 28].

Counterfactual reasoning, or counterfactual thinking (**CT**), is generally defined as a human cognitive ability that captures the process of introspectively reasoning about hypothetical past events, specifically what would have happened if those events had occurred [9, 19]. More simply, counterfactual reasoning, which deals with *what-ifs*, is considered a fundamental component of moral behavior and scientific thought. According to Pearl's Causality Theory, it plays a critical role in the development of thinking machines [28]. In its simplest form, CT may be modelled as an initial form of myopic best response rule at the population level, by considering the fitness of the agent in a system configuration that did not, but could have occurred [19]. Given this new mechanism that greatly empowers individuals' reasoning process the question arises of how the presence of individuals resorting to CT (CTs or counterfactuals [9]) affects the evolutionary dynamics of populations.

The impact of counterfactual reasoning in the context of the evolution of cooperation in populations of agents was first analyzed by Pereira et al. [19], where the authors proposed a model and explored the evolutionary dynamics of a small prevalence of counterfactuals within a population of social learners facing collective action. It is shown that CT has a very low impact on dilemmas that do not require coordination whereas it has a great impact on coordination dilemmas such as the Stag-Hunt Game. Specifically, it is suggested that a small group of counterfactuals strongly fosters the overall population cooperation level, where such behaviour would not be expected. Despite these initial developments being strong insights into the world of counterfactual reasoning in populations, they prompt further investigation. Specifically, we lack understanding regarding the impact of increasing the number of sophisticated agents in the population and whether this insight remains valid across all coordination games.

Here we propose an in-depth study on how counterfactuals influence cooperation in hybrid populations comprising both humans and

---

[*] Corresponding Author. Email: antonio.m.fernandes@tecnico.ulisboa.pt

[1] Professor Ana Paiva ceased her activities on this project on April 4th 2024.

agents. To do so, we apply the evolutionary game theory framework to describe well-mixed hybrid populations of social learners and CT agents, organized in different arrangements. We start by describing the model for homogeneous population dynamics in N-person evolutionary games. Previous studies have shown that counterfactual reasoning promotes the coexistence of strategies [5, 19], thus assuming the role of a coexisting dilemma, therefore shattering coordination games' dynamics. Following this trend, it is of interest to challenge the CT dynamics further by studying the impact of counterfactuals under collective coordination dilemmas, such as the famous N-person stag hunt dilemma (NSH) [26, 38, 47]. More specifically, we investigate how CTs affect cooperation dynamics in relation to the characteristics of the social dilemma.

In the second part, we propose a novel model for counterfactual reasoning in heterogeneous populations. We define the dynamics of collective coordination dilemmas in these populations, based on previous works in heterogeneous populations [4, 47, 48]. We then analyze the framework's dynamics for two sub-populations of social learners and counterfactuals under different conditions, specifically, for different population arrangements and different interactivity levels. From this exploration, we aim to propose some initial insights into this new very complex dynamic.

2 Homogeneous Populations

Let us consider the interaction between a group of N individuals that we assume to be rational and intelligent [22]. These players may either choose to cooperate (C) by contributing a cost c to the public good or to defect (D) by not contributing. After all individuals are given a chance to contribute, the accumulated endowment is multiplied by an enhancement factor F and the amount is equally shared among all individuals of the group. This framework defines a typical public good game, specifically a N-person version of the Prisoner's Dilemma [26]. However, in coordination problems, we assume the public good is only reached if at least $0 \leq M \leq N$ players have contributed, otherwise, no one gets anything. This defines the model for the N-person stag hunt game (NSH), a prototypical example of the social contract [18, 38]. In Table 1 we summarize the payoffs of both C and D, respectively $\Pi_C(k)$ and $\Pi_D(k)$, for any number k of contributors.

Payoff obtained	C	D
$1 \leq k < M$	$-c$	0
$k \geq M$	$\frac{kFc}{N} - c$	$\frac{kFc}{N}$

Table 1. Payoffs values for the N-person stag-hunt game. For later reference, we will define a given NSH with the tuple (N, M, F, C).

We now assume a well-mixed finite population of Z individuals that engage in an NSH, where groups of size N are randomly sampled. For both strategies C and D, we define the average fitness, or social success [13], respectively f_C and f_D, as the hypergeometric distribution for a number k of cooperators:

$$f_C(k) = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k-1}{j} \binom{Z-k}{N-j-1} \Pi_C(j+1) \tag{1}$$

$$f_D(k) = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k}{j} \binom{Z-k-1}{N-j-1} \Pi_D(j) \tag{2}$$

where we impose that the binomial coefficients satisfy the boundary condition of $\binom{k}{j} = 0$ if $k \leq 0$.

Having established the baseline model for interactions among agents, let us now detail the evolution process of finite populations under social learning and then under counterfactual thinking. In both cases this can be conveniently done by adopting the birth-death process, combined with the pairwise comparison rule [45, 46], to describe the evolutionary dynamics of Cs and Ds [36].

Starting with the social learning framework, we say that, at each step, an individual A, following a strategy S_A, imitates a randomly selected individual B following strategy $S_B \neq S_A$ with a probability p given by the standard Fermi distribution [35, 36, 45, 49]:

$$p_{SL}^{S_A \to S_B}(k) \equiv \frac{1}{1 + e^{-\beta_{SL}(f_{S_B}(k) - f_{S_A}(k))}} \tag{3}$$

where β_{SL} stands for the intensity of selection and expresses noises associated with the imitation process errors [19, 36]. Given this model, it now becomes trivial to write down the probabilities to increase (T_{SL}^{+}) or decrease (T_{SL}^{-}) the number of cooperators by one:

$$T_{SL}^{\pm}(k) = \frac{k}{Z}\frac{Z-k}{Z-1}\left[1 + e^{\mp\beta_{SL}(f_C(k) - f_D(k))}\right]^{-1} \tag{4}$$

that are commonly defined as transition probabilities.

The quantity corresponding to the right-hand side of the replicator equation specifies the gradient of selection, a measure that defines the most likely outcome of evolution [26, 36, 41]. In finite populations, the gradient of selection for social learning is given by:

$$G_{SL}(k) \equiv T_{SL}^{+}(k) - T_{SL}^{-}(k) \tag{5}$$

which, in the limit of large populations $Z \gg N$, is equivalent to the replicator equation for infinite populations [26].

Regarding counterfactual reasoning, individuals may not assess the present state of the population but rather evaluate the possible alternative moves to what actually took place. Counterfactuals will conjecture what would be the outcome if their previous decision had been different. Similarly to the SL framework, this can be modelled as an incipient form of the myopic best response rule [19], where an agent A following S_A switched its strategy to $S_B \neq S_A$ with a probability:

$$p_{CT}^{S_A \to S_B}(k) \equiv \frac{1}{1 + e^{-\beta_{CT}(f'_{S_B}(k) - f_{S_A}(k))}} \tag{6}$$

where, like SL, we define β^{CT} as the intensity of counterfactual thinking that, again, regulates the accuracy of the imitation process. Here f'_{S_B} defines the fitness of the strategy S_B in a hypothetical situation had the player chosen S_B in his previous move. In practical terms, this means that a cooperator ($S_A = C$) will compare its strategy's fitness ($f_C(k)$) with the fitness of D if he had chosen that strategy ($f_D(k-1)$). Oppositely, a defector ($S_A = D$) will compare its current strategy's fitness ($f_D(k)$) with the fitness of C if he had chosen that ($f_C(k+1)$).

Unlike social learning, counterfactual reasoning occurs introspectively [5, 19], meaning that one would not require a dyadic interaction to revise its strategy. Taking this into consideration, as before, the transition probabilities for CT come naturally as:

$$T_{CT}^{+}(k) = \frac{Z-k}{Z}\left[1 + e^{-\beta_{CT}(f_C(k+1) - f_D(k))}\right]^{-1} \tag{7}$$

$$T_{CT}^{-}(k) = \frac{k}{Z}\left[1 + e^{-\beta_{CT}(f_D(k-1) - f_C(k))}\right]^{-1} \tag{8}$$

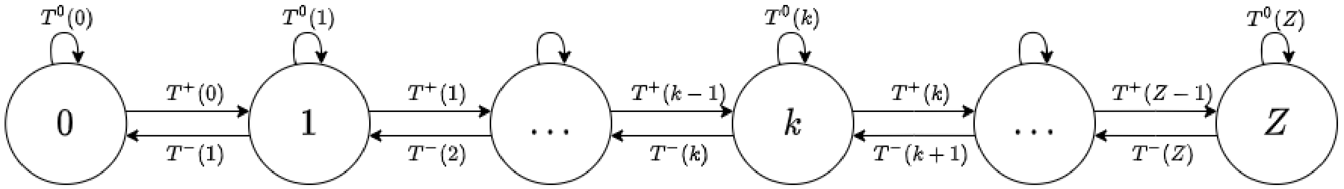

Figure 1. Social dynamics for homogeneous populations as a finite Markov Chain. States are represented by the number k of cooperators in the population. Transition probabilities $T^+(k)$, $T^-(k)$ and $T^0(k)$ represent, respectively, the probabilities to increase, decrease or maintain the number of cooperators by 1.

Notice that, due to the interactions' introspective nature, the CT transition probabilities take a slightly simpler form when compared to SL, not considering the random selection of the second agent in the interaction. This implies a relatively higher order of magnitude when compared to social learning, and hence, for the sake of a better understanding, we are prone to impose $\beta_{CT} > \beta_{SL}$.

In the same way as SL, having these transition probabilities, we define the CT stochastic gradient of selection as:

$$G_{CT}(k) \equiv T^+_{CT}(k) - T^-_{CT}(k) \tag{9}$$

In this framework of homogeneous populations, we assume that individuals are equivalent as they may resort to any of the available heuristics (SL or CT) to revise their strategies. Formally speaking, we define δ as the probability that any individual has to resort to CT in its reasoning (this is equivalent to χ definition in [19]). Naturally, this leads us to a redefinition of the more general system's gradient of selection as:

$$G(k) = (1-\delta)\, G_{SL}(k) + \delta\, G_{CT}(k) \tag{10}$$

Until now, we have not taken into consideration possible mutations in the decision-making process, which raises a big problem in the evolutionary dynamic. The social dynamics framework describes a one-dimensional finite Markov chain with $Z+1$ states, which are represented accordingly to the number k of cooperators [26] (see Figure 1). If we focus on the two possible monomorphic states $k=0$ (or all defectors) and $k=Z$ (all cooperators), we rapidly realize that these states are also absorbing states, meaning that if the system eventually ends up there, it will not leave it. In this case, we say that the update method is non-innovative [37].

To contradict this tendency, however, we can introduce mutations in the definition of the transition probabilities, hence removing the absorbing states and turning the update method to innovative instead. Generally, we say that there is a probability μ of occurring a mutation and an agent adopts a random strategy, without resorting to any of the heuristics SL or CT. As a result, for both SL and CT, we redefine the transition probabilities to:

$$T^+_{SL/CT}(k,\mu) = (1-\mu)\, T^+_{SL/CT}(k) + \mu \frac{Z-k}{Z} \tag{11}$$

$$T^-_{SL/CT}(k,\mu) = (1-\mu)\, T^-_{SL/CT}(k) + \mu \frac{k}{Z} \tag{12}$$

and consequently, we now have:

$$G_{SL/CT}(k,\mu) = T^+_{SL/CT}(k,\mu) - T^-_{SL/CT}(k,\mu) \tag{13}$$

The general gradient of selection then takes its final form homogeneous well-mixed populations as:

$$G(k,\mu) = (1-\delta)\, G_{SL}(k,\mu) + \delta\, G_{CT}(k,\mu) \tag{14}$$

Finally, considering this birth-death process, we can fully describe the Markov process by summarizing all the possible transitions in the transition matrix $\mathcal{T}_{i,j}$ such that $\mathcal{T}_{k,k\pm 1} = T^{\pm}(k,\mu)$ and $\mathcal{T}_{k,k} = 1 - T^-(k,\mu) - T^+(k,\mu)$. From this, we can compute a so-called stationary distribution, $\bar{s}$, which defines the probability of finding the system at a certain state k at any time. It is obtained simply by computing the left eigenvector of the transition matrix with eigenvalue 1, that is, it must satisfy $\bar{s} = \mathcal{T}\bar{s}$ [11, 19].

The gradient of selection and the stationary distribution are the two quantities that will be of foremost importance in the study of population dynamics, as they allow us to predict the most likely outcomes of evolution. To evaluate the cooperation levels under this dynamics, we can define some measures such as the expected fraction of cooperation:

$$EFC = \sum_{k=0}^{Z} k \cdot \bar{s}(k) \tag{15}$$

Additionally, we can compute the overall probability of achieving success, in the form of the group achievement [48], defined as:

$$\eta_G = \sum_{k=0}^{Z} \bar{s}(k) \cdot a_G(k) \tag{16}$$

where we take the multivariate hypergeometric sampling to compute a_G, the (average) fraction of groups that overcome the threshold of M contributors.

2.1 Results and discussion

In line with previous studies, a small prevalence of counterfactuals is enough to push an entire population of social learners towards highly cooperative standards [19]. Despite being very interesting, this insight may raise some questions, specifically, we fall short of knowing what happens if we keep increasing the amount of CT agents in the population and if it is still valid for any coordination game. In this regard, we conduct two studies, in the context of homogeneous populations, that provide helpful insights and a more complete view of the impact of CT in well-mixed populations.

We first compute the general gradient of selection for a homogeneously well-mixed finite population of Z individuals, that play a coordination game, for all the possible values of the probability of CT. As previously stated, to more easily capture the CT impact on the population dynamics, we will assume $\beta_{SL} > \beta_{CT}$. In Figure 2, we show an example of this dynamic for an NSH, with the enhancement factor $M < F < N$, from where two main insights may be taken. The first one, regarding the evolutionary dynamics of the population, is that the CT does promote a coexistence dynamics that completely disrupts and shatters the NSH coordination nature, as expected. If we focus on the evolution of the gradient roots ($G(x, \delta) = 0$), we see that as we increase δ, the NSH stable fixed point on lower x gets closer and closer to the typical unstable fix point [26], eventually cancelling each other and destroying the basin of repulsion.

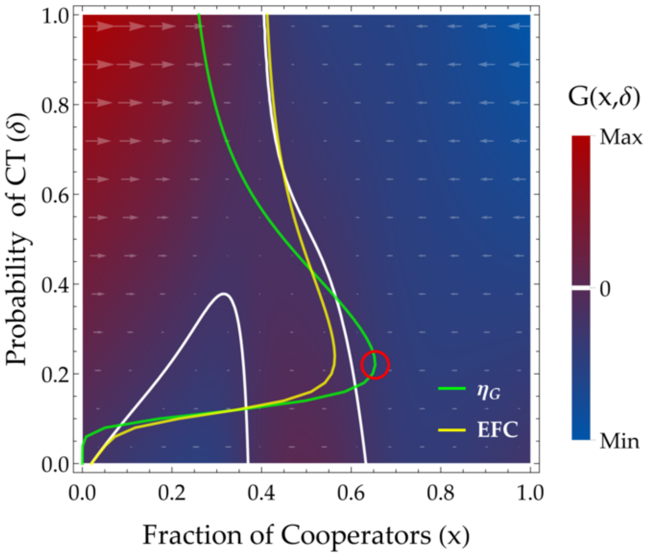

Figure 2. Gradient of selection of a homogeneously well-mixed population as a function of the fraction of cooperators $x = \frac{k}{Z}$ and the probability of CT, δ. We have $Z = 50$ individuals playing a NSH with $(N, M, F, c) = (10, 6, 8.5, 1)$, $\mu = \frac{1}{Z}$ and $\beta_{SL} = 5.0 > \beta_{CT} = 1.0$. Attending to the roots of $G(x, \delta)$, counterfactual reasoning coexistence dynamics completely shatters the coordination nature of the NSH. As we increase the number of counterfactuals, both η_G and EFC rapidly grow until a certain maximum critical point δ_C (represented by a red circle), from which they start to decrease. This indicates there is an optimal amount of counterfactual agents that maximize the cooperation standards.

A second and more interesting insight is that increasing the prevalence of CT-driven agents pushes the entire population towards highly cooperative standards, however, this increase is non-monotonous. As one may infer from both η_G and EFC, increasing the probability of CT leads to an extreme increase of these two quantities until a certain critical point, δ_C, from which they start to decrease, EFC more gently and η_G more abruptly. In other words, although having more counterfactuals in the population is beneficial in terms of cooperation, from a certain δ_C, increasing the probability of CT will progressively prejudice the cooperative standards. This indicates that there is an optimal amount of counterfactual agents that maximize the cooperation standards. Therefore we complement the conclusions previously taken in [19] by stating that a small prevalence of **CT**-driven individuals does promote highly cooperative standards, but only until a critical point, from which having more counterfactuals becomes a disadvantage.

In a second study, we doubt the validity of the previous insights by questioning whether counterfactual reasoning is always beneficial in coordination games. Specifically, we are interested in investigating if there is a dependency on the beneficial action of CT on the coordination game specification. To answer this question, we again take a finite homogeneously mixed population of agents and study the impact of the enhancement factor in the NSH dilemma. We compute the group achievement as $\eta_G(\delta, F)$, that is, as a function of the probability of CT and the enhancement factor, passing through all the different regimes of NSH, $F < M$, $M < F < N$ and $F > N$ [26].

In Figure 3, we show that interestingly $\eta_G(\delta, F)$ demonstrates a different behaviour for each of these three regimes. When $F < M$, CT helps increase the group achievement, hence benefiting cooperation more the higher the δ is. This result comes as trivial if we take into consideration that in this regime cooperators have no chance to survive [26] ($\eta_G \sim 0$) and, as CT promotes coexistence, the only possible outcome of increasing the number of counterfactuals would be to promote cooperation. Oppositely, when $F > N$, that is, in the pure coordination regime [26], CT is (almost) always harmful to cooperation, prejudicing it more the higher the δ is. In this sense, it will be more beneficial to the population to have either a small number or no counterfactuals at all. The only exception resides in a specific configuration where $F \gtrsim N$ and $\delta \sim 0$, where adding a very small amount of counterfactuals in fact nudges the population to very high cooperative standards. However, as we increase δ, cooperative action deteriorates. Finally, the regime where $M < F < N$ is more peculiar, as it shows different behaviours depending if F is closer to M, acting similarly as in the $F < M$ regime, of to N, being closer to the $F > N$ regime. However, in this latter case, there is a big difference: initially, η_G is very small and it rapidly grows, forming this very accentuated slope, that we saw in Figure 2, after which the group achievement begins to decrease, again implying the existence of a critical δ_C for which η_G is maximal. This small region, along with the previously mentioned exception in $F > N$, corresponds to the scenario seen in the previous study, also explored by [19].

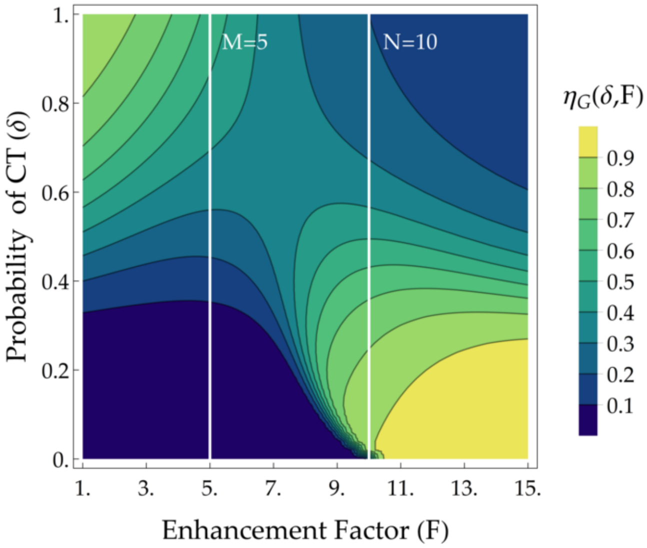

Figure 3. Group achievement as function of the probability of CT, δ and the enhancement factor, F. We consider a finite homogeneous population of $Z = 50$ individuals playing different NSH dilemmas with $(N, M, F, c) = (10, 5, F, 1)$, where $F \in [1, 15]$, $\mu = \frac{1}{Z}$ and $\beta_{SL} = 5.0 > \beta_{CT} = 1.0$. There is a clear distinction between the three regimes: when $F < M$, CT benefits cooperation more the higher the δ is; when $F > N$ CT is mostly harmful for cooperation; when $M < F < N$ we can find a curve where CT aggressively pushes the population towards high cooperative standards for small δ, although it only occurs when F is closer to N than to M. This indicates that the benefit of CT is dependent on the coordination nature of the game, being highly beneficial when $F \lesssim N$.

These results suggest that the benefit of CT is dependent on the coordination nature of the social dilemma. Counterfactual reasoning is generally beneficial to the population in regimes where cooperation is always a disadvantage ($F < N$ [26]), with a more pronounced effect when $F \lesssim N$, and starts to harm the cooperation standards as soon as we enter the pure coordination regime ($F > N$). From a conceptual perspective, the presence of more sophisticated agents in a homogeneous population is more beneficial when cooperation is not expected, and harmful when it is expected.

3 Heterogeneous Populations

When a population is organized in multiple smaller communities of same-heuristic individuals, a different approach shall be taken. Following a framework similar to the one presented in [4, 47, 48], we divide the population according to each individual's learning rule, in our case SL and CT, which is fixed and unchangeable. Here individuals lose the equivalency property that we saw in homogeneous populations. We then treat each sub-population independently, although always considering all interactions between different sub-populations are possible. The interactivity between individuals from different sub-populations is weighted by the homophily, $0 \leq h \leq 1$, that is formally defined as follows: when $h = 0$ anyone in the population may influence and be influenced by anyone else, while when having $h = 1$ individuals are restricted to influence (and be influenced) only by those of the same sub-population. The first case is equivalent to the well-mixed population assumption, while the latter highlights the definition of heterogeneous populations. Homophily is then a measure of the level of heterogeneity, from a well-mixed arrangement ($h = 0$) to a pure heterogeneous one ($h = 1$).

We now consider a population of $Z = Z_{SL} + Z_{CT}$, with Z_{SL} social learners and Z_{CT} counterfactuals, hence one must now consider Cs and Ds for each heuristics. Individuals resorting to each learning rule are given an initial endowment, b_{SL} for SL and b_{CT} for CT agents, which will remain fixed and will play a coordination game, specifically the heterogeneous version of the NSH. Each type of cooperator will contribute to the public good with $c_{SL} = c\,b_{SL}$ and $c_{CT} = c\,b_{CT}$, respectively for SL and CT. For the sake of simplicity, we take $b_{SL} = b_{CT}$. This new dynamic defines a new more complex formulation of the NSH social dilemma, which payoffs can now be defined as:

$$\Pi_C^{SL}(\boldsymbol{k}) = \Pi_D^{SL}(\boldsymbol{k}) - c^{SL} \tag{17}$$

$$\Pi_D^{SL}(\boldsymbol{k}) = (k_{SL} + k_{CT})\frac{Fc\bar{b}}{N}\Theta \tag{18}$$

where, for convenience, we define $\boldsymbol{k} = (k_{SL}, k_{CT})$. The quantity $\Theta \equiv \Theta(c_{SL}k_{SL} + c_{CT}k_{CT} - c\bar{b}M)$ is the Heaviside function, where $\Theta(k) = 1$ whenever $k \geq 0$ and $\Theta(k) = 0$ otherwise. We also define $Z\bar{b} = Z_{SL}b_{SL} + Z_{CT}b_{CT}$ as the average endowment.

The social learner's fitness for C and D, respectively f_C^{SL} and f_D^{SL}, come naturally by using the multivariate hypergeometric sampling without replacement, which according to [47] can be formally presented for social learning as:

$$\begin{aligned} f_C^{SL}(\boldsymbol{k}) =& \binom{Z-1}{N-1}^{-1} \sum_{j_{SL}=0}^{N-1} \sum_{j_{CT}=0}^{N-1-j_{SL}} \binom{k_{SL}-1}{j_{SL}} \binom{k_{CT}}{j_{CT}} \\ & \times \binom{Z - k_{SL} - k_{CT}}{N-1-j_{SL}-j_{CT}} \Pi_C^{SL}(j_{SL}+1, j_{CT}) \end{aligned} \tag{19}$$

$$\begin{aligned} f_D^{SL}(\boldsymbol{k}) =& \binom{Z-1}{N-1}^{-1} \sum_{j_{SL}=0}^{N-1} \sum_{j_{CT}=0}^{N-1-j_{SL}} \binom{k_{SL}}{j_{SL}} \binom{k_{CT}}{j_{CT}} \\ & \times \binom{Z-1-k_{SL}-k_{CT}}{N-1-j_{SL}-j_{CT}} \Pi_D^{SL}(j_{SL}, j_{CT}) \end{aligned} \tag{20}$$

Regarding CT, similarly to the homogeneous populations case, counterfactual reasoning assumes an introspective updating process. Having this, we can write down the fitness for CT as a function of the previous fitnesses as:

$$f_C^{CT}(\boldsymbol{k}) = \begin{cases} f_C^{SL}(k_{SL}, k_{CT}) & C \to D \\ f_C^{SL}(k_{SL}, k_{CT}+1) & D \to C \end{cases} \tag{21}$$

$$f_D^{CT}(\boldsymbol{k}) = \begin{cases} f_D^{SL}(k_{SL}, k_{CT}-1) & C \to D \\ f_D^{SL}(k_{SL}, k_{CT}) & D \to C \end{cases} \tag{22}$$

where $S_A \to S_B$, with $S_{A,B} \in \{C, D\}$, means that an agent following strategy S_A compares its strategy fitness with S_B fitness.

From these fitnesses, we can define the transition probabilities in heterogeneous populations [48] for social learning as:

$$\begin{aligned} T_-^{SL}(\boldsymbol{k}) =& \frac{k_{SL}}{Z}\left\{(1-\mu)\left[\frac{Z_{SL} - k_{SL} + (1-h)(Z_{CT} - k_{CT})}{Z_{SL} - 1 + (1-h)Z_{CT}} \right.\right. \\ & \left.\left. \times \left(1 + e^{-\beta_{SL}\left(f_D^{SL} - f_C^{SL}\right)}\right)^{-1}\right] + \mu\right\} \end{aligned} \tag{23}$$

$$\begin{aligned} T_+^{SL}(\boldsymbol{k}) =& \frac{Z_{SL} - k_{SL}}{Z}\left\{(1-\mu)\left[\frac{k_{SL} + (1-h)k_{CT}}{Z_{SL} - 1 + (1-h)Z_{CT}} \right.\right. \\ & \left.\left. \times \left(1 + e^{-\beta_{SL}\left(f_C^{SL} - f_D^{SL}\right)}\right)^{-1}\right] + \mu\right\} \end{aligned} \tag{24}$$

and for counterfactual thinking as:

$$T_-^{CT}(\boldsymbol{k}) = \frac{k_{CT}}{Z}\left\{(1-\mu)\left[1 + e^{\beta_{CT}(f_C^{CT} - f_D^{CT})}\right]^{-1} + \mu\right\} \tag{25}$$

$$T_+^{CT}(\boldsymbol{k}) = \frac{Z_{CT} - k_{CT}}{Z}\left\{(1-\mu)\left[1 + e^{\beta_{CT}(f_C^{CT} - f_D^{CT})}\right] + \mu\right\}^{-1} \tag{26}$$

where, for the sake of simplicity, we defined the same mutation rate regardless of the heuristic $\mu = \mu_{SL} = \mu_{CT}$.

From these transition equations, we may compute the transition matrix, now adapted to a multi-dimensional space. To do so, we index all the possible states with an integer number, for which conversion we define a bijective function V such that $p = V(\boldsymbol{k})$ and $q = V(\boldsymbol{k'})$, and, consequently, $\boldsymbol{k} = V{-}1(p)$ and $\boldsymbol{k'} = V{-}1(q)$. Afterwards, we may write the transition matrix as $\mathcal{T}_{qp} = T_{\boldsymbol{k}\to\boldsymbol{k'}}$, where $T_{\boldsymbol{k}\to\boldsymbol{k'}}$ is the corresponding transition probability from configuration $\boldsymbol{k}$ to $\boldsymbol{k'}$. Having computed $\mathcal{T}$, one may finally obtain the stationary distribution following the usual procedure. To the obtained vector it is applied the inverse of V we may finally obtain the stationary distribution directly as a function of $\boldsymbol{k}$, as $\bar{s} \equiv \bar{s}(\boldsymbol{k})$ [48].

Also from the transition probabilities, we may generally define the 2-dimensional heterogeneous gradient of selection for social learning and conformity as follows:

$$\nabla(\boldsymbol{k}) = (G_{SL}(\boldsymbol{k}), G_{CT}(\boldsymbol{k})) \tag{27}$$

In this new setup, the group achievement must also be redefined:

$$\eta_G = \sum_{j_{SL}=0}^{Z_{SL}} \sum_{j_{CT}=0}^{Z_{CT}} a_G(\boldsymbol{j}) \cdot \bar{s}(\boldsymbol{j}) \tag{28}$$

with $\boldsymbol{j} = \{j_{SL}, j_{CT}\}$. Additionally, it will also be important to compute the EFC, but now, having different sub-populations, we compute it concerning each of the heuristics as:

$$EFC_L = \frac{k_L}{Z_L} \cdot \bar{s} \tag{29}$$

for each learning rule $L \in \{SL, CT\}$.

3.1 Results and discussion

We have proposed a novel framework for counterfactual thinking in heterogeneous populations, based on previous works of Vasconcelos et al. [47, 48]. This new dynamic counts with two main aspects that strongly differentiate it from the homogeneous framework, described in the previous section. First, we assume individuals are no longer equivalent and their heuristic now comes as a new property for every agent, meaning that the probability of CT now loses its significance. By separating the population heuristics-wise, we now open a new degree of freedom with the proportion between the sizes of each sub-population, Z_{SL}/Z_{CT}. Second, and also as a consequence of this segregation, a new degree of complexity arises as we introduce the concept of homophily, as a measure of the interactivity between the sub-populations. In this respect, we conduct two main studies to investigate the impact of these core differences between the heterogeneous and the homogeneous frameworks.

In the first study, we analyze the different metrics to evaluate the cooperative overall performance, such as η_G and the EFCs for different population compositions. Specifically, by fixing a certain size Z and varying the sizes of both sub-populations, we are able to capture a range of configurations that will allow us to generally evaluate the impact that the proportion between the two communities has on cooperation. In Figure 4 we show the evolution of η_G, EFC, EFC_{SL} and the EFC_{CT} as we increase the proportion between the social learners and counterfactuals, that is, Z_{SL}/Z_{CT}, for different values of the homophily.

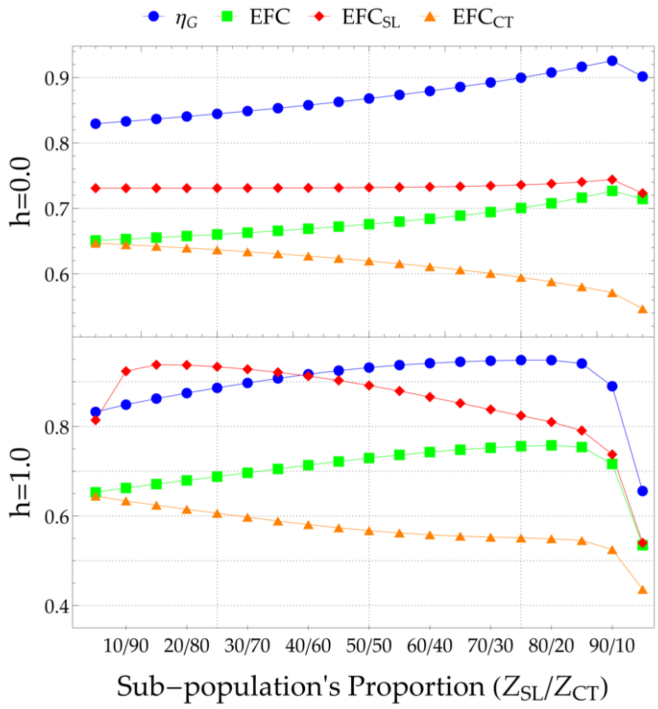

Figure 4. Different cooperation evaluation metrics as a function of the sub-populations proportion, Z_{SL}/Z_{CT}. We fix a population of $Z = 100$ individuals that play an NSH with $(N, M, b_{SL}, b_{CT}, c, F) = (10, 5, 1, 1, 0.3, 10)$, with $\mu = \frac{1}{Z}$ and $\beta_{SL} = \beta_{CT} = 5$. Regardless of the homophily, cooperation reaches a maximum expected value when we have more social learners than counterfactuals. Interestingly, the EFC_{CT} shows to be the opposite, reaching its maximum value the fewer SL agents we have.

Two main insights can be taken from this study. Primarily, irrespective of the homophily, both η_G and EFC assume their maximum values for higher proportions of Z_{SL}/Z_{CT}. Interestingly, EFC_{CT} shows the exact opposite behaviour, assuming its maximum when the number of social learners is minimum. This can suggest that while having fewer counterfactuals promotes the population's overall cooperation to higher standards, the cooperation levels within the CT community decrease. This effect is more pronounced when $h = 0$, as for the opposite case even the EFC_{SL} follow the same trend (as EFC_{CT}).

Secondly, while on the well-mixed populations' regime ($h = 0$) the dynamics seem very stable, with small variations for all cooperation measures when the heterogeneity level is maximal ($h = 1$) a very big slope occurs as we approximate to a very small Z_{CT}. A similar effect occurs in the homogeneous population's dynamics, as demonstrated in Figure 3 when $F \sim N$, where η_G starts close to zero for small δ and grows rapidly as we increase the number of counterfactuals. Similarly, we achieve the same result if we increase Z_{CT} when compared to Z_{SL}. This may be due to the fact that having a very small number of CT-driven agents, we get closer to a single community of SL individuals, that is completely segregated to the CT sub-populations as the homophily is maximal.

In a second study, we investigate the general dynamics of the NSH from a heterogeneous point of view. Particularly, we analyze the gradient of selection and the stationary distribution in a fixed population disposition (specifically having $Z_{SL} = Z_{CT}$) and we vary the level of heterogeneity. In Figure 5 we see an example of this study, where it is shown that a higher homophily slightly increases the group achievement, hence promoting better cooperative outcomes. Another (much more) interesting insight is that the maximum value of the stationary distribution progressively moves towards higher values of k_{SL}, although slightly moving downward toward lower values of k_{CT}. This indicates that the most likely outcome of evolution will end up in highly cooperative standards among social learners but not among counterfactual thinkers, a result that is in agreement with the previous study in which we saw a decreasing EFC_{CT} as we increase the proportion Z_{SL}/Z_{CT}.

Although very intriguing, notice that these conclusions are extremely dependent on the game dynamics and therefore one should not generalize for any coordination dilemma. As we previously did for the homogeneous populations, a more complete study is necessary to complement the study of this complex heterogeneous model for counterfactual reasoning.

4 Conclusions

In this work, we unravel the impact of counterfactual reasoning in the decision-making process under collective action, specifically coordination dilemmas, for different population arrangements. In the context of well-mixed finite populations, we started by defining the evolutionary framework when individuals are considered equivalent, hence able to adopt any heuristic for their decision process, thus establishing a homogeneously mixed population. In this regard, we complement the conclusions from Pereira et al. [19] by showing that a small prevalence of CT-driven individuals may indeed promote highly cooperative standards, but only until a critical point, from which having more counterfactual thinkers becomes a disadvantage. Moreover, we show that, in the case of an NSH, this result is only valid when the enhancement factor assumes values close to the group size ($F \lesssim N$), hence suggesting that the CT impact is dependent on the social game characteristics. We saw very distinct

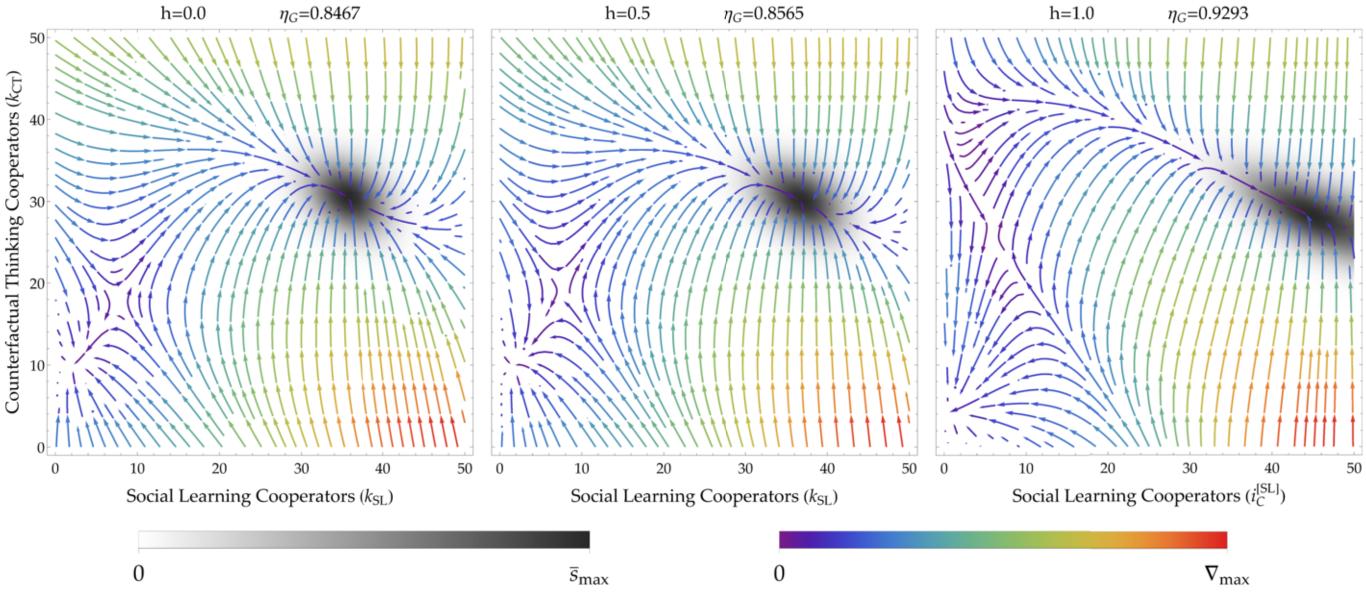

Figure 5. Gradient of selection and stationary distribution of a heterogeneous population playing a coordination game, for different homophilies $h = \{0, 0.5, 1\}$. Individuals are playing an NSH described by $(N, M, b_{SL}, b_{CT}, c, F) = (10, 5, 1, 1, 0.3, 9.5)$. We have $Z_{SL} = Z_{CT} = 50$, with $\mu = \frac{1}{Z}$ and $\beta_{SL} = \beta_{CT} = 5$. We notice a small growth in the group achievement as we increase the homophily, generally suggesting a positive effect of CT. Additionally, the stationary distribution shows that an increase in the heterogeneity level pushes the system towards high cooperative standards among social learners, while slightly harming CT cooperative action.

impacts of CT on the NSH, being generally beneficial when cooperation is not expected ($F < M$), and mostly harmful when it is expected ($F > N$). This suggests that the presence of individuals employing counterfactual reasoning may not necessarily benefit the population as a whole.

Furthermore, we proposed an evolutionary model for the population's heterogeneous arrangement, where we now consider that individuals are not equivalent and assume a specific heuristic at all times, hence establishing different communities, or sub-populations, for each learning rule. Compared with the homogeneous framework, despite losing the definition of the probability of CT (δ), this new dynamic now considers two new degrees of freedom: the communities' sizes proportion (Z_{SL}/Z_{CT}) and the interactivity, measured through the homophily. Regarding the first measure, we showed that the group achievement and the EFC are maximal for a higher relative amount of SLs (large Z_{SL}/Z_{CT}), a result that goes in agreement with the previous study on a homogeneous population. However, interestingly the opposite occurs among counterfactuals, as EFC_{CT} decreases with increasing Z_{SL}/Z_{CT}. This result suggests that while having fewer CT-driven individuals boosts the population's overall cooperation to higher standards, the cooperation levels within the CT community decrease. This insight was later reinforced in the last study, where we analyzed the gradient of selection and the stationary distribution for an NSH, now varying the homophily. Specifically, we showed that an increase in homophily not only leads to a slight enhancement of the group achievement but also pushes the stationary distribution towards high cooperative standards within social learners, despite gently lowering the expected cooperation levels for counterfactual thinking. Notice, however, that these studies were made in the specific case of a N-person coordination game in the regime of $F \lesssim N$, hence providing a very incomplete panorama of the heterogeneous population dynamics. In this regard, we suggest more complete studies for different populations by exploring different social dilemmas and metrics with the proposed model as a baseline.

An understanding of the reasoning behind the decision-making process is essential to leverage cooperation in futuristic hybrid societies of humans and machines [5, 27, 19, 28]. To better comprehend how humans behave, either naturally or facing artificial entities, it is required a complete perception of the how human mind works either individually or from a community perspective. The modulation of the human mind, from a population point of view, passes through the understanding of different heuristics such as natural selection (through social learning) [1, 22, 40, 39], conformity [17, 43, 44], counterfactual thinking [17, 19, 28], dissimilarity [10], aspiration-based learning [3, 51], among others. To add another level of complexity, human reasoning is not as basic as the theoretical game theory framework proposes, as individuals frequently engage in deeper decision mechanisms that take into account the beliefs and goals of others to possibly predict future actions. These non-trivial thinking mechanisms can be usually modelled through the lens of Theory of Mind [6, 32, 52]. This additional layer of complexity enables much richer dynamics, offering ample room for further exploration [16, 52].

Having such complexity, we suggest that a proper general formalization of evolutionary game theory, which considers all mentioned heuristics and dynamics, is required to help in the unification of all different frameworks explored so far. This is, again, imperative if we want to accurately build a complete model of human behaviour and the human mind.

Acknowledgements

This research was supported by the INESC-ID (UIDB/50021/2020), as well as by the project CRAI C645008882-00000055/510852254 (IAPMEI/PRR). We appreciate constructive feedback from the anonymous reviewers for improving the final version of this paper.

References

[1] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *science*, 211(4489):1390–1396, 1981.

[2] C.-H. Chan, H. Yin, P. Hui, and D.-F. Zheng. Evolution of cooperation in well-mixed n-person snowdrift games. *Physica A: Statistical Mechanics and its Applications*, 387(12):2919–2925, 2008.

[3] F. Chen, L. Zhou, and L. Wang. Cooperation among unequal players with aspiration-driven learning. *Journal of the Royal Society Interface*, 21(212):20230723, 2024.

[4] M. C. Couto, J. M. Pacheco, and F. C. Santos. Governance of risky public goods under graduated punishment. *Journal of Theoretical Biology*, 505:110423, 2020.

[5] M. C. Couto, S. Giaimo, and C. Hilbe. Introspection dynamics: a simple model of counterfactual learning in asymmetric games. *New Journal of Physics*, 24(6):063010, 2022.

[6] H. de Weerd, D. Diepgrond, and R. Verbrugge. Estimating the use of higher-order theory of mind using computational agents. *The BE Journal of Theoretical Economics*, 18(2):20160184, 2018.

[7] M. Doebeli and C. Hauert. Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology letters*, 8(7):748–766, 2005.

[8] W.-B. Du, X.-B. Cao, L. Zhao, and M.-B. Hu. Evolutionary games on scale-free networks with a preferential selection mechanism. *Physica A: Statistical Mechanics and its Applications*, 388(20):4509–4514, 2009.

[9] K. Epstude and N. J. Roese. The functional theory of counterfactual thinking. *Personality and social psychology review*, 12(2):168–192, 2008.

[10] Y. Fang, T. P. Benko, M. Perc, and H. Xu. Dissimilarity-driven behavior and cooperation in the spatial public goods game. *Scientific Reports*, 9 (1):7655, 2019.

[11] A. R. Góis, F. P. Santos, J. M. Pacheco, and F. C. Santos. Reward and punishment in climate change dilemmas. *Scientific reports*, 9(1):16193, 2019.

[12] C. S. Gokhale and A. Traulsen. Evolutionary games in the multiverse. *Proceedings of the National Academy of Sciences*, 107(12):5500–5504, 2010.

[13] T. A. Han, L. M. Pereira, and F. C. Santos. The emergence of commitments and cooperation. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '12, page 559–566, Richland, SC, 2012. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 0981738117.

[14] M. Hu, W. Chen, M. Yang, and Q. Li. Evolution of cooperation based on synergy and discounting on n-person snowdrift game. In *2022 41st Chinese Control Conference (CCC)*, pages 6922–6926. IEEE, 2022.

[15] S. Kurokawa and Y. Ihara. Emergence of cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, 276 (1660):1379–1384, 2009.

[16] T. Lenaerts, M. Saponara, J. M. Pacheco, and F. C. Santos. Evolution of a theory of mind. *Iscience*, 27(2), 2024.

[17] J. Lin, C. Huang, Q. Dai, and J. Yang. Evolutionary game dynamics of combining the payoff-driven and conformity-driven update rules. *Chaos, Solitons & Fractals*, 140:110146, 2020.

[18] Q. Luo, L. Liu, and X. Chen. Evolutionary dynamics of cooperation in the n-person stag hunt game. *Physica D: Nonlinear Phenomena*, 424: 132943, 2021. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2021.132943. URL https://www.sciencedirect.com/science/article/pii/S0167278921001019.

[19] L. Moniz Pereira and F. C. Santos. Counterfactual thinking in cooperation dynamics. *arXiv e-prints*, pages arXiv–1912, 2019.

[20] J. Moreira, J. Vukov, C. Sousa, F. C. Santos, A. F. d'Almeida, M. D. Santos, and J. M. Pacheco. Individual memory and the emergence of cooperation. *Animal Behaviour*, 85(1):233–239, 2013.

[21] L. Mui, M. Mohtashemi, and A. Halberstadt. Notions of reputation in multi-agents systems: a review. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pages 280–287, 2002.

[22] R. B. MYERSON. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991. ISBN 9780674341166. URL http://www.jstor.org/stable/j.ctvjsf522.

[23] M. A. Nowak. *Evolutionary dynamics: exploring the equations of life*. Harvard university press, 2006.

[24] M. A. Nowak. Five rules for the evolution of cooperation. *science*, 314 (5805):1560–1563, 2006.

[25] H. Ohtsuki and Y. Iwasa. How should we define goodness?—reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, 231 (1):107–120, 2004. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2004.06.005. URL https://www.sciencedirect.com/science/article/pii/S0022519304002772.

[26] J. M. Pacheco, F. C. Santos, M. O. Souza, and B. Skyrms. Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proceedings of the Royal Society B: Biological Sciences*, 276(1655):315–321, 2009.

[27] A. Paiva, F. Santos, and F. Santos. Engineering pro-sociality with autonomous agents. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[28] J. Pearl and D. Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, Inc., USA, 1st edition, 2018. ISBN 046509760X.

[29] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki. Statistical physics of human cooperation. *Physics Reports*, 687: 1–51, 2017.

[30] J. Pi, G. Yang, and H. Yang. Evolutionary dynamics of cooperation in n-person snowdrift games with peer punishment and individual disguise. *Physica A: Statistical Mechanics and its Applications*, 592: 126839, 2022.

[31] D. G. Rand and M. A. Nowak. Human cooperation. *Trends in cognitive sciences*, 17(8):413–425, 2013.

[32] T. Rusch, S. Steixner-Kumar, P. Doshi, M. Spezio, and J. Gläscher. Theory of mind and decision science: Towards a typology of tasks and computational models. *Neuropsychologia*, 146:107488, 2020.

[33] F. C. Santos and J. M. Pacheco. Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences*, 108(26):10421–10425, 2011.

[34] F. C. Santos, J. M. Pacheco, and T. Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences*, 103(9):3490–3494, 2006.

[35] F. P. Santos, S. Encarnação, F. C. Santos, J. Portugali, and J. M. Pacheco. An evolutionary game theoretic approach to multi-sector coordination and self-organization. *Entropy*, 18(4):152, 2016.

[36] M. D. Santos, F. L. Pinheiro, F. C. Santos, and J. M. Pacheco. Dynamics of n-person snowdrift games in structured populations. *Journal of Theoretical Biology*, 315:81–86, 2012.

[37] M. D. d. A. Santos. Evolutionary dynamics of collective action in structured populations. 2012.

[38] B. Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.

[39] J. M. Smith. Evolution and the theory of games. In *Did Darwin get it right? Essays on games, sex and evolution*, pages 202–215. Springer, 1982.

[40] J. M. Smith and G. R. Price. The logic of animal conflict. *Nature*, 246 (5427):15–18, 1973.

[41] M. O. Souza, J. M. Pacheco, and F. C. Santos. Evolution of cooperation under n-person snowdrift games. *Journal of Theoretical Biology*, 260 (4):581–588, 2009.

[42] Q. Su, A. Li, L. Wang, and H. Eugene Stanley. Spatial reciprocity in the evolution of cooperation. *Proceedings of the Royal Society B*, 286 (1900):20190041, 2019.

[43] A. Szolnoki and M. Perc. Conformity enhances network reciprocity in evolutionary social dilemmas. *Journal of The Royal Society Interface*, 12(103):20141299, 2015.

[44] A. Szolnoki and M. Perc. Leaders should not be conformists in evolutionary social dilemmas. *Scientific Reports*, 6(1):23633, 2016.

[45] A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic dynamics of invasion and fixation. *Physical Review E*, 74(1):011909, 2006.

[46] A. Traulsen, J. M. Pacheco, and M. A. Nowak. Pairwise comparison and selection temperature in evolutionary game dynamics. *J Theor Biol*, 246 (3):522–529, 2007.

[47] V. V. Vasconcelos, F. C. Santos, and J. M. Pacheco. A bottom-up institutional approach to cooperative governance of risky commons. *Nature Climate Change*, 3(9):797–801, 2013.

[48] V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, and S. A. Levin. Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences*, 111(6):2212–2216, 2014.

[49] V. V. Vasconcelos, F. P. Santos, F. C. Santos, and J. M. Pacheco. Stochastic dynamics through hierarchically embedded markov chains. *Phys Rev Lett*, 118(5):058301, 2017.

[50] W.-X. Wang, J. Ren, G. Chen, and B.-H. Wang. Memory-based snowdrift game on networks. *Physical Review E*, 74(5):056113, 2006.

[51] X. Wang, C. Gu, J. Zhao, and J. Quan. Evolutionary game dynamics of combining the imitation and aspiration-driven update rules. *Physical Review E*, 100(2):022411, 2019.

[52] W. Yoshida, R. J. Dolan, and K. J. Friston. Game theory of mind. *PLoS computational biology*, 4(12):e1000254, 2008.