Ecological Modelling 439 (2021) 109348

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A multi-armed bandit algorithm speeds up the evolution of cooperation

Roberto Cazzolla Gatti [a,b]

[a] Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria
[b] Biological Institute, Tomsk State University, Tomsk, Russian Federation

ARTICLE INFO

Keywords:
Evolution of cooperation
Multi-armed bandit algorithm
Epsilon-greedy model
Matryoshka model

ABSTRACT

Most evolutionary biologists consider selfishness an intrinsic feature of our genes and as the best choice in social situations. During the last years, prolific research has been conducted on the mechanisms that can allow cooperation to emerge "in a world of defectors" to become an evolutionarily stable strategy. A big debate started with the proposal by W.D. Hamilton of "kin selection" in terms of cost sustained by the cooperators and benefits received by related conspecifics. After this, four other main rules for the evolution of cooperation have been suggested. However, one of the main problems of these five rules is the assumption that the payoffs obtained by either cooperating or defeating are quite well known by the parties before they interact and do not change during the time or after repeated encounters. This is not always the case in real life. Following each rule blindly, there is a risk for individuals to get stuck in an unfavorable situation. Axelrod (1984) highlighted that the main problem is how to obtain benefits from cooperation without passing through several trials and errors, which are slow and painful. With a better understanding of this process, individuals can use their foresight to speed up the evolution of cooperation. Here I show that a multi-armed bandit (MAB) model, a classic problem in decision sciences, is naturally employed by individuals to opt for the best choice most of the time, accelerating the evolution of the altruistic behavior and solving the abovementioned problems. A common MAB model that applies extremely well to the evolution of cooperation is the epsilon-greedy (ε-greedy) algorithm. This algorithm, after an initial period of exploration (which can be considered as biological history), greedily exploits the best option ε% of the time and explores other options the remaining percentage of times (1-ε%). Through the epsilon-greedy decision-making algorithm, cooperation evolves as a multilevel process nested in the hierarchical levels that exist among the five rules for the evolution of cooperation. This reinforcement learning, a subtype of artificial intelligence, with trials and errors, provides a powerful tool to better understand and even probabilistically quantify the chances cooperation has to evolve in a specific situation.

1. Introduction

If the "struggle for life" is what moves Darwinian evolution on (Darwin, 1859), why those species that cooperate more are some of the most successful and widespread on our planet? Social insects (such as ants, wasps, termites, and bees), school of fishes, flocks of birds, herd of herbivores, pack of wolves, group of primates, human families, cities and even institutions all offer good evidence that competition might not be the most important natural selector.

Though Nature has long been considered "red in tooth and claw" (Fry 2012), such as to justify the competitive behavior of our species from cultural, biological, and economic perspectives (Gould 1992; Hull 1999; Marriott 2002; Numbers and Numbers 2003), altruism may be more advantageous in an individual's lifespan and lead to the long term evolution of cooperation (Nowak et al., 2000; Gatti 2011).

Several studies proposed that, rather than continually fighting, wasting energy and time, species can evolve mutualistic relationships, which over time makes the avoidance of competition more convenient than the competition itself and can emerge in symbiotic (Kiers and Heijden 2006), endosymbiotic (O'Malley 2015) and endogenosymbiotic (Cazzolla Gatti 2016, 2018) interspecific relations. These mutualistic networks stabilize the ecosystems (Bastolla et al., 2009; Cazzolla Gatti et al. 2017, 2018) and improve their resilience and productivity (Thébault and Fontaine 2010). However, mutualism, as well as inter- and intra-specific cooperation can be stable and prosperous but can be undermined by cheaters or selfish behaviours. Nonetheless, the benefits of intraspecific cooperation (the altruistic help provided by one to an individual of its same species) remain a controversial issue in evolutionary biology, because it is thought that being selfish is an intrinsic feature of our genes (Dawkins 2016) and most times it is the best choice

---

E-mail address: robertocgatti@gmail.com.

https://doi.org/10.1016/j.ecolmodel.2020.109348
Received 22 September 2020; Received in revised form 19 October 2020; Accepted 20 October 2020

in social situations (Midgley 1983). Evolution would favor egoistic individuals increasing their fitness (i.e. the number of their offspring or their reproductive success; Jost 2003). In fact, it seems undeniable that, in simplified non-zero-sum games (such as the Prisoner's dilemma; Lave 1962) where players individually interact with each other just once or for a fixed number of times, defeating cooperation seems the best choice (an Evolutionary Stable Strategy, ESS, i.e. a Nash equilibrium; Hofbauer 2001) although not the most advantageous (in terms of maximum reachable payoff). During the last fifty years, prolific research has been conducted on the mechanisms that can allow cooperation to emerge "in a world of defectors" to become an ESS. An interesting debate started with the simple but genial "kin selection" proposed by W. D. Hamilton (1964), which - following the rule named after its author - can be beneficial for both the donor and the recipient of an altruistic act (an action costly to perform but benefitting another individual of the same species) if they are enough related. For instance, two siblings would have more interest to cooperate than two strangers because they share a higher percentage (about 50% more) of their genes. Successively, based on Hamilton's reasoning in terms of cost sustained by the cooperators and benefits received by conspecifics, the other four main rules for the evolution of cooperation have been suggested (Nowak 2006). However, one of the main problems of these five rules is the assumption that the payoffs obtained by either cooperating or defeating are quite well known by the parties before they interact and do not change during the time or after repeated encounters (Nowak et al., 2017). Obviously, this is not always the case in real life. Moreover, this does not assume that all individuals of a population have to benefit from a cooperative strategy. The strategy may only be beneficial at the population level over evolutionary timescales. Any optimization at the individual level may counteract this, which indicates that individuals may not optimally explore the consequence of their behavior to obtain an evolutionarily successful strategy.

At the same time, following each rule blindly, there is a risk for individuals to get stuck in an unfavorable situation. Nonetheless, this is not necessarily universal. Altruism may not always provide an advantage for the individual donor. Therefore decisions may just happen by instinct and not because these decisions are actively judged every time. The main problem is how to obtain benefits from cooperation without passing through several trials and errors, which are slow and painful (Axelrod, 1984). This would also increase the costs (investment of time, energy, etc.) over the benefits of altruistic behavior. In fact, even if the conditions for cooperation are convenient in the long term, individuals may not have sufficient time to explore the strategies to obtain a stable reward from their altruism. In The Evolution of Cooperation, Axelrod (1984) wrote that "if we understand the process we can use our foresight to speed up the evolution of cooperation". Here I show that a multi-armed bandit (MAB) model, a classic problem in decision sciences, is naturally employed by individuals to opt for the best choice most of the time following the hierarchical nestedness of the rules, accelerating the evolution of the altruistic behavior and solving the abovementioned problems. The proposed algorithm takes also into account situations of instinctive altruism that omits any active judgment that many cooperating individuals may not be able to practice.

2. Background: the 5 rules for the evolution of cooperation, their limits and optimization

The basic rules for the evolution of cooperation (kin selection and direct reciprocity) are accepted by many evolutionary biologists and seem widespread among almost all groups of living beings (Rothstein and Pierotti 1988; Zahavi 1995). Nevertheless, it is obvious that, at least, some members of our species do not act altruistically only with relatives or people that expect to meet in the future. We often offer help to unrelated people that we, surely, meet once in a life. Do we behave in this way because we are just a "good citizen"? In terms of Darwinian fitness, this does not make any sense because we waste our energy, time, resources, etc. (sustaining a cost that reduces our own fitness) to provide support (donate a benefit that increases another person's fitness) to a competitor (in a classic evolutionary definition). Being altruists with a stranger must give us an advantage, natural selection states. Here is where other forms of cooperation show off (Table 1).

All the forms of cooperation described in Table 1 could represent a widespread interaction model even for some lower taxa of animals (invertebrates), plants, fungi, and microorganisms (see Table 2). Nonetheless, some authors have argued that the 5 rules for the evolution of cooperation are actually not all distinct from each other and represent a continuum in the evolutionary pattern toward cooperation (Lehmann and Rousset 2014; West et al., 2007).

It is easy to notice that all these five rules are related by the a priori estimation of costs and benefits and that, to be evolutionary stable, each parameter (r, w, q, 1/k, $1/(1+n/m)$) must exceed the c/b ratio. The costs must be lower than the benefits to have an evolutionary advantage. Of course, if the costs are higher than benefits, none of the five rules can stabilize cooperation because, for instance, no relatedness value can exceed the unit (i.e. meeting with a perfect clone/twin) or no probability higher than 100% to meet another individual again.

Quantify benefits and costs, particularly in the case of kin selection, is not an easy task because these parameters are related to the variation in average trait value and therefore cannot predict that change (Nowak et al., 2017). For instance, Hamilton's rule seems able to estimate only the data that have already been provided and this makes impossible any reliable empirical test. Because social interactions are usually multilateral and nonlinear, it is difficult to constrain them with unique benefit and cost values (Nowak et al., 2017). Complex population structures cannot be captured by a single relatedness quantity.

In reciprocal altruism, nonetheless, Tit-for-Tat is evolutionarily stable only if the interactions between the individuals have enough probability to continue in the future (Axelrod and Dion, 1988). However, Always-Defect is also evolutionary stable even with a low probability of future interactions (Imhof et al., 2005). Therefore, how cooperative behavior can start at the beginning? There must be a first step that allows some probabilities of altruistic acts before moving to reciprocity. Genetic kinship seems to be a likely way to bypass the defection equilibrium. Then the reciprocal forms of cooperation can take the place of kin selection.

Reciprocity (both direct and indirect), however, can stabilize in species with the ability to discriminate against other conspecifics. But when the group is too big and variable or there are cognitive impediments for group member's recognition, cooperation can be still stable. Other mechanisms, which do not require a high level of discrimination, include territoriality with set location, network establishment, grouping and gregarize (Akçay et al., 2010; Grinnell 2002). These lead altruism towards network and group selection.

Besides the difficulties in estimating the cost and benefits of an altruistic act and the instability of a single specific cooperation mechanism, there is the risk for individuals - following blindly each rule - to stuck into an unfavorable situation (for instance, continuing to cooperate) if the conditions change over time. Aging or illness of the partner might change the convenience of an altruistic act, decreasing (or even sometimes increasing: e.g., in the case of keeping in the grace of the old member of a group to receive its legacy) the propensity to cooperate (Szolnoki et al., 2009).

Optimizing the strategies that fit best in each situation is another relevant issue, which has been investigated in several other behavioral contexts such as feeding optimization (see, for instance, Vadas et al., 1994 and Lupatsch et al., 2001), cooperative foraging optimization (see, for instance, Chen et al., 2009), avoidance of predation optimization (see, for instance, Cazzolla Gatti et al. 2020), travel optimization (see, for instance, Lihoreau et al., 2010), habitat choice optimization (see, for instance, Schmitz 1991), reproductive optimization (see, for instance, Oksanen and Lundberg 1995), etc. For a general review on optimization theory in evolution see Smith (1978) and Rosenhouse (2002).

Table 1
The five rules for the evolution of cooperation briefly described.

Rule	Equation	Parameter	Description
KIN SELECTION	$r>c/b$	***r***	Cooperation can become the evolutionarily stable strategy (ESS, in game theory the one in which a population of cooperators cannot be invaded by defectors) in the case of kin selection (1st rule; Hamilton 1964) if the relatedness between the two interacting individuals is higher than the costs/benefits ratio ($r>c/b$).
DIRECT RECIPROCITY	$w>c/b$	***w***	Altruistic behavior could not only be promoted by relatedness. For instance, relatedness is not always enough to explain altruistic behaviors such as eusociality (Nowak and Allen 2015). In conditions of repeated meetings (repeated Prisoner's dilemma) a sort of direct reciprocity (2nd rule) can be established (Trivers 1971). This is based on each other acquaintance after repeated interactions. Evolution of cooperation can take place (following either a Tit-for-Tat, a generous-Tit-for-Tat or a win-stay/lose-shift strategy; Nowak and Sigmund 1992, 1993) if the probability of next encounter between the two individuals (w) exceeds the c/b ratio ($w>c/b$).
INDIRECT RECIPROCITY	$q>c/b$	***q***	Individuals may behave altruistically with an unrelated, once-in-a-life met, another individual, if someone is looking at me and, being cooperative, I can increase my reputation. I help you today because someone else, knowing that I am an altruist, will help me in turn (Nowak and Sigmund, 1998). This implies that the parties should have a "theory of mind" and be aware that similar strategies can be displayed by different players (i.e. what is good for me can be good for you too, later on). Although this form of "reputational altruism" can be challenged by free-riders (someone knows that I am an altruist and will take advantage of this by defeating cooperation) and requires some sorts of moral assessments (will my reputation increase or decrease, in the eyes of other people, if I either help or not someone that has not previously been cooperative? Nowak and Sigmund 2005), it becomes evolutionary stable (3rd rule) if the probability (q) of knowing somebody's reputation is higher than the c/b ratio ($q>c/b$).
NETWORK RECIPROCITY	$1/k>c/b$	1/k	We cannot always assume that repeated meetings happen in well-mixed populations where the probability to interact with everyone else is equal (Leimar and Hammerstein 2001a, 2001b) because this is not the real situation in many cases. Natural populations can be gregarious, spatially and/or socially structured and some individuals can meet more likely than others. This is the reason why network reciprocity has been advanced as a mechanism (4th rule) to explain how cooperation can evolve in aggregated populations (Ohtsuki and Nowak 2007) in which the number of neighbors (k) may affect my propensity for altruistic acts. To be evolutionary stable, network reciprocity requires that the inverse of the number of neighbors (1/k) exceeds the cost/benefit ratio ($1/k>c/b$).
GROUP SELECTION	$1/(1+n/m)>c/b$	$1/(1+n/m)$	As the evolution of sociality and sociobiology clearly suggests, natural selection can act not only at an individual level but also at a group level (5th rule), which means that aggregations of cooperators can defeat other less or non-cooperative groups (Traulsen and Nowak 2006). Imagine a population of individuals of the same species that is subdivided into some groups of members that help each other more frequently (cooperators) and others that do not behave altruistically (defectors). Cooperation can stabilize and evolve at a group level if the inverse of $1+n/m$ (where n is the maximum group size and m is the number of groups) is higher than c/b ratio ($1/(1+n/m)>c/b$).

3. The model

3.1. A multi-armed bandit to speed up the evolution of cooperation

Consider the following situation: you meet a person in a train station who needs help to find his/her train but you have just a few minutes to get on your own train. Your help could cost you to wait more than two hours for the next train. Would you, however, help this person? How do you decide what to do in a few seconds? The theory says that if this person is enough related to you (he is your brother, she is your cousin, etc.) you might afford a cost to provide a benefit to him/her if $r>c/b$. This could be a quick choice, assuming if you know the grade of relatedness with this person. However, it is not so easy to correctly evaluate in a hurry situation whether your w (chance to meet this person again), q (someone is paying attention to your act), etc. exceeds the costs to benefits ratio. Human beings have brains but not computers. We have a "theory of mind" but cannot perform an accurate calculus in a train station. We can even find it difficult to evaluate the real costs and benefits of our actions (i.e. to build a payoff matrix as in game theory) and keep track of their changes during time. Other species might not have all these reasonings behind their altruistic acts (most have the skills to generate the benefits but a few are recognized to have those for distributing them; Calcott 2008; Markvoort et al., 2014). In most cases, animals may find it difficult even to identify the level of relatedness with other members of their populations (Schweinfurth and Taborsky 2018; Stiver et al., 2005). Therefore, how can individuals choose the best option more frequently? Anticipating Hamilton's rule (1964), it is narrated that once Haldane said: "I would gladly give up my life for two brothers or eight cousins" (Gorrell et al., 2010). Yet, this does not help in coupled interactions with a single brother, cousin, etc. if $b>c$ does not compensate for the fact that $r<1$ or b and c are difficult to estimate. Therefore, how species can speed up the evolution of cooperation invoked by Axelrod, which leads to higher payoff for everyone when it stabilizes?

The multi-armed bandit (MAB), a classic problem in decision sciences (Even-Dar et al., 2006), could help opt for the best choice most of the time, accelerating the evolution of the altruistic behavior and solving the previous paradoxes. Its name origins from old slot machines where to win money you need to pull an arm (called "bandit" because they steal money from the player). If you have, in front of you, two (or more) machines and you know that their payoff (i.e. how much money you get pulling a certain arm) is not equal between the machines, your best option is to identify the machine with the highest payoff. Once you have this guess you start to pull this machine more than the other one. This is a problem of optimal resource allocation under uncertainty: how do I most efficiently identify the best machine to play, while efficiently exploring the other option in real-time, so I can exploit the best an get the highest payoff?

This approach has the power to bring the theoretical assumptions described so far to the everyday situation in which an individual has to decide whether to cooperate or not with someone else. The basic model to decide which strategy to select in an unknown or poorly known strategy payoff is to perform repeated A/B tests, which can be time consuming and inefficient. If the choice on the model takes too much time, the response could be delayed and useless in the optimal decision-making process. In the example of the train station, if I took too long to decide whether to help the person needing advice, both trains could leave and both will pay only costs, without receiving benefits. Similarly, if I accelerate my decision and I base it on pure chance or a few A/B test scores, I have higher probabilities to take the wrong choice and my costs as a donor may outperform the benefits for the recipient. Moreover, even in the simplest event of kin selection, when I am interacting with a sibling, there could be changing situations (aging or illness, for instance), those in which I cannot clearly evaluate costs and benefits and whether their ratio is lower than 50%, etc. Should I miss my train to help my brother to take his one? Although, cooperation can evolve by natural selection without a conscious understanding of the evolutionary rationale of it, for individuals of species that lack a "theory of mind" the rational evaluation of the pros and cons of an action is impossible. Thus, there should be an easy, quick and reliable way living beings take –

Table 2
Evidence of different forms of intraspecific cooperation in Nature from selected literature (question marks denote unclear/partial evidence, stars denote potential/unproved existence, and slashes denote likely impossibility of a specific form of cooperation).

	r Kin selection	w Direct reciprocity	q Indirect reciprocity	k Network reciprocity	$1 + n/m$ Group selection
Microorganisms	West et al., 2006, Wingreen and Levin 2006, Griffin et al., 2004, Leeks and West 2019	*	/	*	?, Swenson et al., 2000
Plants	Dudley and File 2007, File et al., 2012, Karban and Shiojiri 2009, Murphy and Dudley, 2009	?, Graham and Bormann, 1966, Fajardo and McIntire, 2011, Baret and DesRochers, 2011	/	?, Robinson and Fitter 1999, Walder et al., 2012	?, Graham and Bormann, 1966 Fajardo and McIntire, 2011, Baret and DesRochers, 2011
Invertebrates (other than social insects)	Duffy et al., 2002, Duffy and Macdonald, 2010	*	/	*	?, de Jager et al. 2020
Social insects	Queller and Strassmann 1998, Kennedy and Radford 2020, Mathiron et al., 2019	?, Nee 1989, Freidin et al., 2017, Carter 2014	/	*	Nowak et al., 2010, Birch 2019, Wilson and Hölldobler 2005
Vertebrates (other than Primates and Hominidae)	Lee 1987, Griffin and West 2003, Russell and Hatchwell 2001	Carter and Wilkinson 2013, Olendorf et al., 2004, Brandl and Bellwood 2015, Wilkinson, 1988	?, Akçay et al., 2010, Doutrelant and Covas 2007, Truskanov et al., 2020	?, Croft et al., 2005, Truskanov et al., 2020	?, Muir 1996, Goodnight and Craig 1996
Primates (non-Hominidae)	Silk 2002, Eberle and Kappeler 2006, Chapais et al., 2001	Seyfarth and Cheney 1984, De Waal and Brosnan 2006, Molesti and Majolo 2017	?, Majolo et al., 2012, Anderson et al., 2017, Schweinfurth and Call 2019	?, Schino and Aureli 2010	?, Boehm 1981, Aoki and Nozawa 1984
Hominidae	Morin et al., 1994, Silk 2006, Rand and Nowak 2013	Jaeggi and Gurven, 2013, Jaeggi et al., 2010, Kopp and Liebal 2016	Nowak and Sigmund 2005, Leimar and Hammerstein 2001a, 2001b, Subiaul et al., 2008	Yamamoto and Tanaka 2009, Wang et al., 2013, Szolnoki and Perc 2015	Wilson and Wilson 2008, Rachlin 2019

conscious or unconscious - decisions to increase the likelihood of adopting the right strategy.

3.2. The epsilon-greedy algorithm for the evolution of cooperation

Evolution is a matter of chances and necessities (Monod 1974) and it always seeks the simplest way to explore solutions and exploit the best ones (Carroll 2001). Furthermore, adaptation is a matter of time, of fast rearrangement of strategies considering that the best choice now might not be the optimal one in the future.

Bandit algorithms are models to real-time and real-life decision making that represent an efficient trade-off between sufficiently exploring the best options during the time and exploiting the optimal action in current situations. This is the kind of trade-off favored by natural selection when striking a balance between the two is critically important. Firstly, the variant space needs to be sufficiently explored such that the best choice is identified. By first identifying, then continuing to exploit the optimal action, it is possible to maximize the total payoff that is available to an individual from the environment. However, an individual also wants to continue to explore other feasible variants in case they provide better returns in the future (if some conditions will change). That is, you want to avoid committing yourself when faced with a difficult choice somewhat by continuing to explore (a little) sub-optimal variants if the related payoff changes. For instance, becoming older you might not have any doubt whether to help your sister at the train station if she has to be in time at a job interview and you are a plenty-of-time retired person, even if this means that you will have to wait two hours for the next train. Nonetheless, you could get no gratitude, no reward and no reputation from your altruism and the result would be that, being late, you just missed your favorite TV show while your sister is having professional egoistic success. Similarly, for a social bee, altruism towards an aging queen could become less advantageous in terms of scarifying workers' reproduction (De Loof 2011). Thus, the choice to be selfish is still on the plate. What to do, then? If the conditions change adapting to new situations, the bandit algorithm will consider this variation and will begin selecting a new strategic behavior for a new meeting. During the exploration of the costs and benefits, the interacting individuals (even if unconsciously) learn more about the average payoff rate and what is the distribution of uncertainty around that (Akçay and Roughgarden 2011).

The evolutionary key (which could make a real difference on an individual's fitness, i.e. gene survival) during the decision process is, therefore, to assess how best to balance the exploration-exploitation tradeoff. A common MAB model, which applies extremely well to the evolution of cooperation, to balance this tradeoff is the epsilon-greedy ($\varepsilon-$greedy) algorithm (Sutton and Barto 1998).

This algorithm, after an initial period of exploration (biological history, which can either be considered as evolutionary history in the case of kin, network and group selection or life history in the case of shorter-term adaptive interactions, such as direct or indirect reciprocity), greedily exploits the best option ε percent of the time. For example, if ε =0.95, the algorithm will exploit the best variant 95% of the time and will explore random alternatives 1- ε=5% of the time. Depending on the environment (which might be the full set of resources and conditions available at that specific time) an individual runs a bandit algorithm in its mind (enough automatic, easy and fast to do not require a "high cognitive level" and an estimation of costs and benefits) to decide what strategy to take. It then performs the action, observes the payoff and updates its knowledge of the reward. With a sort of reinforcement learning (RL) (Kalidindi and Bowman 2007) evolution would lead each individual to get a basic knowledge of its payoff while interacting with another party and to take advantage of this information to exploit, during its life, the best choice more frequently (ε% of times) but leave a bit of space to exploration (1- ε%) to adapt to the changing conditions.

A formal description of an epsilon-greedy algorithm is the following:

```
- Initialize, for $a = 1$ to k
- Q(a) ← 0
- N(a) ← 0
- loop for n times
- A ¬ $\begin{cases} argmax_a(Q_a) \rightarrow with\ probability\ \varepsilon \\ a\ random\ action \rightarrow with\ probability\ 1-\varepsilon \end{cases}$
- R ← bandit(A)
- N(A) ← N(A) + 1
- Q(A) ← $Q(A) + \frac{1}{N(A)}[R - Q(A)]$
```

where a is a random number of k actions, Q(a) is the estimate of expected reward for taking the action a, N(a) is the number of time action a is selected, R is the set of Rewards obtained, and A is the set of Actions taken.

A slightly modified version of the epsilon-greedy algorithm for the evolution of cooperation (Fig. 1) can, therefore, be described as the following:

- randomly choose Cooperation (C) or Defection (D) for n trials (biological history);
- on each trial estimate the payoff of variant C and D;
- set the parameters r, w, q, 1/k, $1/(1 + n/\mathrm{m}) = \varepsilon$;
- after n learning trials (biological history):
  - if C payoff $> D$ payoff:
    - (a) select (exploit) ε% of the time the variant C;
    - (b) 1−ε% of the time sample (explore) from the variants C and D randomly;
  - when D payoff $> C$ payoff:
    - (a) switch the choice (σ) and select (exploit) 1-ε% of the time the variant D;
    - (b) ε% of the time sample (explore) from the variants D and C randomly.

It is important to point out that this is a slightly modified version of the original epsilon-greedy algorithm. Here a switch (σ) is included. This is because the parameters r, w, q, 1/k, $1/(1 + n/\mathrm{m})$ are the epsilon value. Thus ε% represents the number of times cooperation should be exploited when its payoff is higher than that of selfishness. Obviously, the higher the parameter value ε, the more convenient to exploit cooperation ε% of the time. Therefore, the epsilon value is predetermined by evolutionary constraints and depends on the biological history of relationships and interactions among interacting individuals. If, however, during interactions the best payoff "switches to" selfishness, the individual starts to explore both variants ε% of the time (Fig. 1). In this case, in fact, the lower the parameter value ε, the more convenient to exploit selfishness 1- ε% of the time. This switch easily corrects for evaluation mistakes and bias in payoff estimations after the learning trials (for instance, even if r is ≥ 0.5, there might be situations when D payoff is evaluated higher than C payoff and this would lead to overexploitation of relatives without a switch function σ).

The initial exploration trials of the epsilon-greedy algorithm follow an Evolutionary-Development (Evo-Devo) pattern. They are performed by some individuals during evolutionary history (Evo) and selected for the best fitness. This constitutes pre-birth information genetically transferred, as in the case of kin and group selection. But they can be also performed during the life history of every single individual (Devo) that could be advantaged in its fitness. This constitutes learning from life experiences, as in the case of direct, indirect and network reciprocity. In both cases, these trials provide an estimate of the initial payoff for cooperation and defection about the parameters r, w, q, etc.

Following a hierarchical decision-making process of the five rules in sequence (Fig. 2), individuals select ε% (i.e. $= r$%, w%, q%, etc.) of the time to exploit the payoff rate gained from cooperation. However, an individual will select 1-ε% of the time to sample from two variants. Either cooperation can be selected (1-ε%)/2 of times or to move to the lower cooperative rule can be selected with the same probability (1-ε%)/2, randomly (like just flipping an imaginary coin) to leave a bit of space to exploration to adapt to the changing conditions in future. If, after the hierarchical exploration of all the five rules (or of only those species-specific that apply, see Table 1), cooperation is not achieved, then the individual will defect and behave selfishly. After each interaction, however, through reinforcement learning, each individual will acquire more knowledge of its payoff for potential next interaction with the same individual or group. If the payoff for selfishness becomes higher than that assumed for cooperation, the individual switches (σ) to exploit selfishness 1- ε% of the time and explore randomly between the two variants ε/2% of the time.

In the five circumstances depicted in Fig. 2 it emerges that **ε+(1-ε)/2**

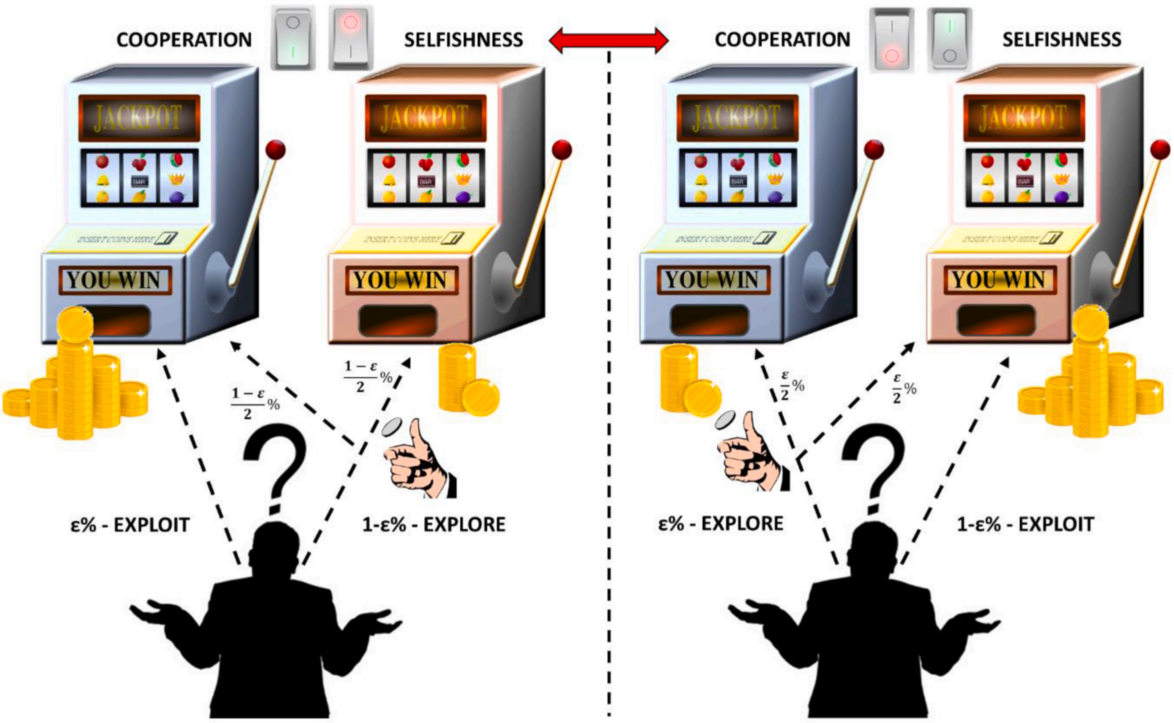

Fig. 1. The modified epsilon-greedy algorithm for the evolution of cooperation is based on a multi-armed bandit model. Each slot machine returns a payoff and the player aims to maximize it. If the player learns from biological history that the "cooperation" slot machine has a higher payoff than the "selfish" slot machine (left panel), it will exploit (ε% of the time) cooperation and explore randomly the two options (1- ε% of the time). However, during the time the best payoff could switch from cooperation to selfishness (right panel) and the player will exploit (1-ε% of the time) selfishness and explore randomly the two options (ε% of the time).

$\boldsymbol{> (1\text{-}\varepsilon)/2}$ for any $\varepsilon = r$, w, q, 1/k and $1/(1 + n/\text{m})$. Therefore, when C payoff $> D$ payoff, cooperation quickly evolves with a minimum probability $P_{ECmin} = \varepsilon + (1\text{-}\varepsilon)/2$ whereas selfishness has a maximum probability to evolve of $P_{ESmax} = (1\text{-}\varepsilon)/2$.

For instance:

- if $r = 0.5$ (two siblings interact), $\varepsilon = 0.5$ and $0.5 + 0.25 > 0.25 = 0.75 > 0.25$, the probability of cooperation to evolve is minimum 75% and the probability of exploration for the alternative variant (lower hierarchical cooperation level and, eventually, selfishness) is maximum 25%
- if $w = 0.25$ (two interacting individuals have 1 out of 4 chances to meet again), $\varepsilon = 0.25$ and $0.25 + 0.375 > 0.375 = 0.575 > 0.375$, the probability of cooperation to evolve is minimum 57.7% and the probability of exploration for the alternative variant is maximum 37.5%
- if $q = 0.125$ (the probability of knowing someone's reputation in presence of 7 other unknown individuals), $\varepsilon = 0.125$ and $0.125 + 0.4375 > 0.4375 = 0.5625 > 0.4375$, the probability of cooperation to evolve is minimum 56.25% and the probability of exploration for the alternative variant is maximum 43.75%

It is trivial to prove that, for any value of r, w, q, 1/k and $1/(1 + n/\text{m})$, P_{EC} is always > 0.5. Hence, cooperation has always more chances to evolve than competition following one of the five rules, when this epsilon-greedy algorithm is applied in decision making and the payoff for cooperation is the highest, skipping the need to estimate costs and benefits of an interaction. Nonetheless, besides speeding up the evolution of cooperation, the algorithm still keeps some space open to evolving conditions with a $(1\text{-}\varepsilon)/2$ probability to explore other choices (selfishness, in this case, and improve the future payoff).

4. Discussion

This MAB model encompasses and provides a comprehensive framework for the five rules of cooperation. Therefore, the model is not an alternative to the five rules but a logic structure for optimizing them. The five rules for the evolution of cooperation, through a multi-armed bandit choice selection, will still hold without keeping an individual in an inconvenient situation if it cannot exactly estimate the costs and benefits and/or the environmental changes (including variations in individual behavior, network and group compositions, etc.) during the time.

An evident hierarchical ranking exists among the five rules for the evolution of cooperation (Fig. 2), with kin selection at the first level, followed by direct, indirect, and network reciprocity, and then by group selection. This means that (1- r)/2% of times an individual, before opting to be a selfish (i.e. defeat cooperation), moves its decision-making process to the lower hierarchical level and evaluates the choice based on the second parameter (w=probability of next round). Again, (1-w)/2% of the times the decision shifts to the next level and so on, until it reaches the last one (group selection) where $(1\text{–}1/(1 + n/\text{m}))/2\%$ of times selfishness is the only likely option. This hierarchical shifting of cooperation rules can be summarized by the expression:

$$\text{r} \rightarrow \text{q} \rightarrow \text{w} \rightarrow 1/\text{k} \rightarrow 1/(1+n/\text{m}) \rightarrow \varepsilon\% \text{ cooperation} | 1 - \varepsilon\% \text{ exploration}$$

From this whole scenario it emerges that cooperation, through the epsilon-greedy decision-making algorithm, evolves as a multilayer process: horizontally (algorithmically) and vertically (hierarchically).

It is interesting to notice that in an evolutionary scenario of cooperation based on this multi-armed bandit model, haplodiploidy results in one the highest probability of cooperation in Nature. This is the sex-determination system of Hymenoptera and other taxa of invertebrates, in which males develop from unfertilized eggs and are haploid, and females develop from fertilized eggs and are diploid. Therefore insects with a coefficient r of relatedness of 0.75, which is common between the sisterhood of worker bees in a hive, have $r = \varepsilon = 0.75$, $P_{ECmin} = 0.875$ and $P_{ESmax} = 0.125$. However, while haplodiploid females are more closely related to their sisters than to their offspring, they still share more genes with their offspring than with their brothers ($r = 0.25$). Previous research suggests that the evolutionary burden of raising low-value brothers would, therefore, offset the advantages of rearing high-value sisters. Nonetheless, the epsilon-greedy algorithm resolves this imbalance already at the first level (kin selection) because when $r = \varepsilon = 0.25$, P_{ECmin} is still 65% and P_{ESmax} is 25%. Therefore, cooperating with brothers is still worth it in a cooperative hive (i.e. where initial trials during biological history assessed that C payoff is generally higher than D payoff).

Overall, from the hierarchical epsilon-greedy algorithm, we can predict that cooperation has the following general probability to evolve when C payoff $> D$ payoff:

$$P_{EC} = r * \text{w} * \text{q} * 1/\text{k} * 1/(1 + n/\text{m})$$

In the previous example of a beehive with a sisterhood of worker bees' relatedness (r) among sisters is 0.75 and the probability of a next interaction (w) between two of them is extremely high (let's assume = 0.9). Considering that there is no evidence of indirect reciprocity for invertebrates (see Table 1), but network and group selections have been suggested, at least for social insects (Fewell 2003), cooperation among social bees evolves with a probability $= 0.75 \times 0.9 \times 1/\text{k}*1/(1 + n/\text{m})$. Which is: $0.675 \times 1/\text{k}(1 + n/\text{m})$. This means that, besides relatedness and direct reciprocity, the evolution of eusociality is strongly dependent on the number of neighbors (k) and the size/number of groups $(1 + n/\text{m})$. This confirms the recent evidence (Field & Toyoizumi 2020) from Hymenoptera with more specialized morphological castes that two common features of eusociality (i.e. saturating birth rates and group size) significantly determine whether eusociality outperforms other strategies.

4. Conclusions

How well Hamilton's rule can account for all the different forms of altruism is a long-debated topic in evolutionary biology. The key point relates to the levels of selection: kin selection favors more cooperation among related individuals and, on the other extreme, group selection enlarges the interactions within and between entire groups of organisms. So far, the problem resided in the idea that selection between groups cannot be stable enough to promote adaptations because selection acts mostly at the individual and not group level. Classical evolutionary biology considered between-group a very special case that cannot account for most of the cooperative interactions. Though, recently, some evidence that kin selection and multilevel selection can be mathematically equivalent has emerged, the continuum of different cooperation levels and its hierarchical interconnection has been largely ignored.

The relation between heritable traits and fitness has been mostly based on direct versus indirect (for kin selection) and within groups versus between groups (for group selection) benefits. Hamilton's rule and group selection, instead, should be considered just as the starting and ending point of multilevel interactions. Examining the evolution of cooperation on actual individuals, biologists should follow a Matryoshka model proposed by Wilson and Wilson (2008).

Because natural selection operates at multiple levels of the biological hierarchy, the Matryoshka model, combined with the epsilon-greedy algorithm, should be used to evaluate the importance of each level on a case-by-case basis (Fig. 3).

As shown above, the interacting layers of the evolution of cooperation are like a matryoshka: nested one within another. At each hierarchical level natural selection favors a different rule of cooperation. The general rule is that cooperation at a certain level requires a corresponding process of selection at that level and might be weakened by selection at lower levels. Case-by-case it should be shown how a

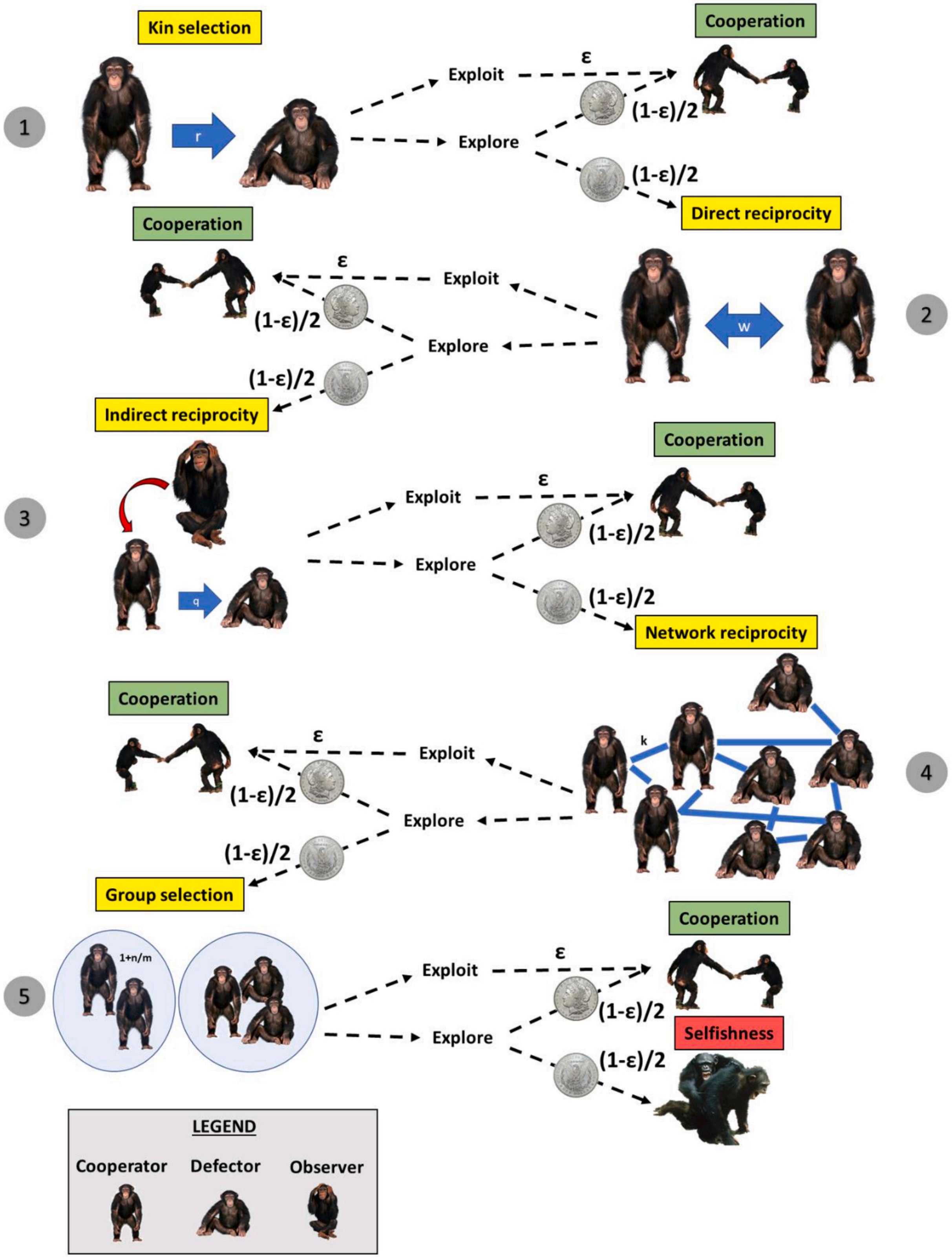

Fig. 2. The epsilon-greedy algorithm applied to the hierarchical decision choice through the 5 cooperation rules, when the payoff for cooperation exceeds that of selfishness. Starting from level 1 (kin selection) the individual chooses to cooperate ε% (where ε=*r*) of the time and (1-ε)% of the time flips a coin (i.e. randomly choose) to explore the two variants. Therefore, (1-ε)/2% of the time the choice moves to the lower level. With the same probability, the process continues in this way until it reaches the last level (group selection) where the individual opts ε%+(1-ε)/2% of the time for cooperation and (1-ε)/2% for selfishness.

cooperative level of selection is strong enough to counterbalance the higher selective advantage for selfishness. An interesting example comes from a model organism mentioned by Wilson and Wilson (2008).

The amoebae *Dictyostelium discoides* is an important species for the study of evolutionary development (Evo-Devo). In these organisms, the stalk cells in the multicellular slug seem to sacrifice their lives “for the good of the group” following the principles of multilevel selection. In fact, the spore cells in the head of the fruiting body successfully disperse and pass their genes to the next generation only if other cells form a stalk and sacrifice their own reproduction. But why some altruistic cells should give up to improve their own fitness to give an advantage to other selfish cells? It might be that “altruistic cells” recognize themselves and aggregate only with others that carry the same altruistic genes so that each stalk is genetically similar. In this case, however, natural selection would favor cooperation at the group level but altruistic genes (in the stalk) would be reduced up to almost disappear after a certain number of selfish generations (in the fruiting body). A more plausible explanation is now offered by the multilevel selection (matryoshka model) applied to the epsilon-greedy algorithm. As in a lottery process, some individuals make the ultimate sacrifice to become stalk cells but they are a random sample deriving from the ε exploration of cooperative and selfish strategies. In this way, there is no genetic selection within a group and the altruistic genes have a higher chance to survive at the next generations both to the benefit of the individual carrying cells and whole multicellular organisms.

To better investigate the evolution of cooperation we must, therefore, look first to kin selection but, if this does not seem the likely reason of an altruistic act, we should move forward, at a higher level, to include other ecological, evolutionary, and developmental (Eco-Evo-Devo) variables relevant to explain a specific interaction through the epsilon-greedy algorithm.

The empirical analysis of the evolution of cooperation should follow a five-question scheme:

1) What is the percentage of the individual’s relatedness with its partner?
2) What is the probability it has to meet its partner again?
3) What is the chance that someone is looking at its interaction with its partner?
4) What is the number of cooperative neighbors in its social network?
5) What is the number of cooperative members and groups in its environment?

This approach may be the basis for further, more application-oriented, studies. Through reinforcement learning, a subtype of artificial intelligence based on the idea that a computer learns as a living being does, with trials and errors, we would have a powerful tool to better understand and even probabilistically quantify the chances cooperation has to evolve in a specific situation. This would solve most of the problems of estimating a priori costs and benefits of an altruistic act and adapt to the dynamical change of the external conditions. For instance, a behavioral ecologist studying the interactions among members of a group or individuals of different species may use a computer simulation to understand the dynamics and the evolution of decision making and predict the probability of cooperation evolution by: i) gathering information about the biological history of the individuals/species involved to check if the payoff was higher in previous cooperative or competitive interactions (i.e. whether the “switch” is set on payoff $C>D$ or payoff $D>C$); ii) setting the parameters (r, w, q, k, $1+n/m$) that apply to the subject involved in the study; iii) using the

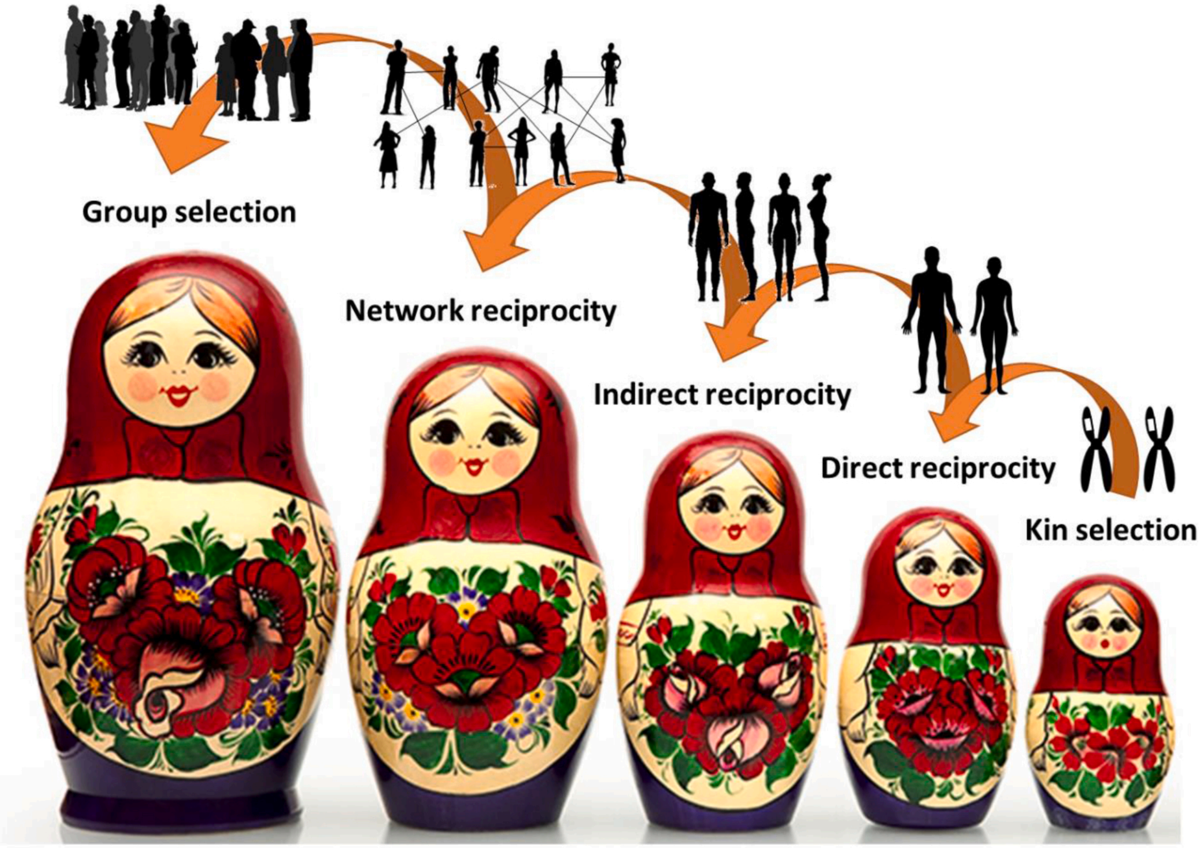

Fig. 3. To study the evolution of cooperation biologists should follow a Matryoshka ε-greedy model in which kin selection is the smallest doll (lowest level) and group selection the biggest one (highest level). Between them, the other levels are nested hierarchically and allow to explain the likelihood of altruist behaviors by individuals.

Matryoshka ε-greedy algorithm proposed here to estimate the $\varepsilon\%+(1-\varepsilon)/2\%$ at each step of the nested hierarchical roles; iv) running the model to look at the dynamical variation of the cooperative/selfish behavior when the parameters and the payoffs change during a pre-established time (e.g. number of encounters).

This multi-armed bandit model to the evolution of cooperation answers the question of how should an individual dedicate a defined amount of resources to different options when it cannot be sure what will come of (i.e. which payoff) pulling each one. With the epsilon-greedy algorithm, as life and evolutionary time goes on, and the individual is choosing different options, it will get a sense of which choices are returning it with the highest reward. However, from time to time it will choose a random action just to make sure that it's not missing anything. Using this learning algorithm, the individual converges to the optimal strategy in choosing between cooperation and selfishness.

Credit author

Roberto Cazzolla Gatti is the only author of this paper

Declaration of Competing Interest

The author declare no conflict of interest

References

Akçay, Ç., Reed, V.A., Campbell, S.E., Templeton, C.N., Beecher, M.D., 2010. Indirect reciprocity: song sparrows distrust aggressive neighbours based on eavesdropping. Anim Behav 80 (6), 1041–1047.

Akçay, E., Roughgarden, J., 2011. The evolution of payoff matrices: providing incentives to cooperate. Proc. Biol. Sci. 278 (1715), 2198–2206.

Anderson, J.R., Bucher, B., Chijiiwa, H., Kuroshima, H., Takimoto, A., Fujita, K., 2017. Third-party social evaluations of humans by monkeys and dogs. Neurosci. Biobehav. Rev. 82, 95–109.

Aoki, K., Nozawa, K., 1984. Average coefficient of relationship within troops of the Japanese monkey and other primate species with reference to the possibility of group selection. Primates 25 (2), 171–184.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Axelrod, R., Dion, D., 1988. The further evolution of cooperation. Science 242 (4884), 1385–1390.

Baret, M., DesRochers, A., 2011. Root connections can trigger physiological responses to defoliation in nondefoliated aspen suckers. Botany 89 (11), 753–761.

Bastolla, U., Fortuna, M.A., Pascual-García, A., Ferrera, A., Luque, B., Bascompte, J., 2009. The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature 458 (7241), 1018–1020.

Birch, J., 2019. Are kin and group selection rivals or friends? Curr. Biol. 29 (11). R433-R438.

Boehm, C., 1981. Parasitic selection and group selection: a study of conflict interference in rhesus and Japanese macaque monkeys. Primate Behavior and Sociobiology. Springer, Berlin, Heidelberg, pp. 161–182.

Brandl, S.J., Bellwood, D.R., 2015. Coordinated vigilance provides evidence for direct reciprocity in coral reef fishes. Sci Rep 5, 14556.

Calcott, B., 2008. The other cooperation problem: generating benefit. Biol Philos 23 (2), 179–203.

Carroll, S.B., 2001. Chance and necessity: the evolution of morphological complexity and diversity. Nature 409 (6823), 1102–1109.

Carter, G., 2014. The reciprocity controversy. Anim Behav Cogn 1 (3), 368–386.

Carter, G.G., Wilkinson, G.S., 2013. Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment. Proc. Biol. Sci. 280 (1753), 20122573.

Cazzolla Gatti, R., 2018. Endogenosymbiosis: from hypothesis to empirical evidence towards a Unified Symbiogenetic Theory (UST). Theor Biol Forum 111 (1–2), 13–26.

Cazzolla Gatti, R., Fath, B., Hordijk, W, Kauffman, .S, Ulanowicz, R., 2018. Niche emergence as an autocatalytic process in the evolution of ecosystems. J. Theor. Biol. 454, 110–117.

Cazzolla Gatti, R., Hordijk, W., Kauffman, S., 2017. Biodiversity is autocatalytic. Ecol. Model. 346, 70–76.

Cazzolla Gatti, R., Messina, G., Tiralongo, F., Ursino, L.A., Lombardo, B.M., 2020. Learning from the environment: how predation changes the behavior of terrestrial Isopoda (Crustacea Oniscidea). Ethol. Ecol. Evol. 32 (1), 29–45.

Chapais, B., Savard, L., Gauthier, C., 2001. Kin selection and the distribution of altruism in relation to degree of kinship in Japanese macaques (Macaca fuscata). Behav. Ecol. Sociobiol. (Print) 49 (6), 493–502.

Chen, H., Zhu, Y., Hu, K., 2009. Cooperative bacterial foraging optimization. Discrete Dyn. Nat. Soc. 2009.

Croft, D.P., James, R., Ward, A.J.W., Botham, M.S., Mawsdley, D., Krause, J., 2005. Assortative interactions and social networks in fish. Oecologia 143 (2), 211–219.

Darwin, C., 1859. On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life. John Murray, London.

Dawkins, R., 2016. The Selfish Gene. Oxford university press.

de Jager, M., van de Koppel, J., Weerman, E.J., Weissing, F.J., 2020. Patterning in mussel beds explained by the interplay of multi-level selection and spatial self-organization. Front. Ecol. Evol. 8, 7.

De Loof, A., 2011. Longevity and aging in insects: is reproduction costly; cheap; beneficial or irrelevant? A critical evaluation of the "trade-off" concept. J. Insect Physiol. 57 (1), 1–11.

De Waal, F.B., Brosnan, S.F., 2006. Simple and complex reciprocity in primates. Cooperation in Primates and Humans. Springer, Berlin, Heidelberg, pp. 85–105.

Doutrelant, C., Covas, R., 2007. Helping has signalling characteristics in a cooperatively breeding bird. Anim. Behav. 74 (4), 739–747.

Dudley, S.A., File, A.L., 2007. Kin recognition in an annual plant. Biol. Lett. 3 (4), 435–438.

Duffy, E.J., Morrison, C.L., Macdonald, K.S., 2002. Colony defense and behavioral differentiation in the eusocial shrimp Synalpheus regalis. Behav. Ecol. Sociobiol. (Print) 51 (5), 488–495.

Duffy, J.E., Macdonald, K.S., 2010. Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis. Proc. Biol. Sci. 277 (1681), 575–584.

Eberle, M., Kappeler, P.M., 2006. Family insurance: kin selection and cooperative breeding in a solitary primate (Microcebus murinus). Behav. Ecol. Sociobiol. (Print) 60 (4), 582–588.

Even-Dar, E., Mannor, S., Mansour, Y., 2006. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. J. Mach Learn Res 7 (Jun), 1079–1105.

Fajardo, A., McIntire, E.J., 2011. Under strong niche overlap conspecifics do not compete but help each other to survive: facilitation at the intraspecific level. Journal of Ecology 99 (2), 642–650.

Fewell, J.H., 2003. Social insect networks. Science 301 (5641), 1867–1870.

File, A.L., Murphy, G.P., Dudley, S.A., 2012. Fitness consequences of plants growing with siblings: reconciling kin selection, niche partitioning and competitive ability. Proc. Biol. Sci. 279 (1727), 209–218.

Freidin, E., Carballo, F., Bentosela, M., 2017. Direct reciprocity in animals: the roles of bonding and affective processes. Int J. Psychol. 52 (2), 163–170.

Fry, D.P., 2012. Life without war. Science 336 (6083), 879–884.

Gatti, R.C., 2011. Evolution is a cooperative process: the biodiversity-related niches differentiation theory (BNDT) can explain why. Theor. Biol. Forum 104 (1), 35–43.

Gatti, R.C., 2016. A conceptual model of new hypothesis on the evolution of biodiversity. Biologia 71 (3), 343–351.

Goodnight, C.J., Craig, D.M., 1996. The effect of coexistence on competitive outcome in Tribolium castaneum and Tribolium confusum. Evolution (N Y) 50 (3), 1241–1250.

Gorrell, J.C., McAdam, A.G., Coltman, D.W., Humphries, M.M., Boutin, S., 2010. Adopting kin enhances inclusive fitness in asocial red squirrels. Nat Commun 1 (1), 1–4.

Gould, S.J., 1992. Red in tooth and claw. Nat Hist 101 (11), 14–20.

Graham, B.F., Bormann, F.H., 1966. Natural root grafts. The Botanical Review 32 (3), 255–292.

Griffin, A.S., West, S.A., 2003. Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. Science 302 (5645), 634–636.

Griffin, A.S., West, S.A., Buckling, A., 2004. Cooperation and competition in pathogenic bacteria. Nature 430 (7003), 1024–1027.

Grinnell, J., 2002. Modes of cooperation during territorial defense by African lions. Hum. Nat. 13 (1), 85–104.

Hamilton, W.D., 1964. The genetical evolution of social behaviour. II. J. Theor. Biol. 7 (1), 17–52.

Hofbauer, J., 2001. From Nash and Brown to Maynard Smith: equilibria, dynamics and ESS. Selection 1 (1–3), 81–88.

Hull, D.L., 1999. Evolutionists red in tooth and claw. Nature 398 (6726), 385-385.

Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. Proc. Nat. Acad. Sci. 102 (31), 10797–10800.

Jaeggi, A.V., Gurven, M., 2013. Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: a phylogenetic meta-analysis. Proc. Biol. Sci. 280 (1768), 20131615.

Jaeggi, A.V., Stevens, J.M., Van Schaik, C.P., 2010. Tolerant food sharing and reciprocity is precluded by despotism among bonobos but not chimpanzees. Am. J. Phys. Anthropol. 143 (1), 41–51.

Jost, J., 2003. On the notion of fitness, or: the selfish ancestor. Theor. Biosci. 121 (4), 331–350.

Kalidindi, K., Bowman, H., 2007. Using ϵ-greedy reinforcement learning methods to further understand ventromedial prefrontal patients' deficits on the Iowa Gambling Task. Neural Netw. 20 (6), 676–689.

Karban, R., Shiojiri, K., 2009. Self-recognition affects plant communication and defense. Ecol. Lett. 12 (6), 502–506.

Kennedy, P., Radford, A.N., 2020. Sibling quality and the haplodiploidy hypothesis. Biol. Lett. 16 (3), 20190764.

Kiers, E.T., Heijden, M.G.V.D., 2006. Mutualistic stability in the arbuscular mycorrhizal symbiosis: exploring hypotheses of evolutionary cooperation. Ecology 87 (7), 1627–1636.

Kopp, K.S., Liebal, K., 2016. Here you are!—Selective and active food sharing within and between groups in captive Sumatran orangutans (Pongo abelii). Behav. Ecol. Sociobiol. (Print) 70 (8), 1219–1233.

Lave, L.B., 1962. An empirical approach to the prisoners' dilemma game. Q J Econ 424–436.

Lee, P.C., 1987. Sibships: cooperation and competition among immature vervet monkeys. Primates 28 (1), 47–59.

Leeks, A., West, S., 2019. Altruism in a virus. Nat. Microbiol. 4 (6), 910–911.

Lehmann, L., Rousset, F., 2014. The genetical theory of social behaviour. Philos. Trans. R. Soc. B Biol. Sci. 369 (1642), 20130357.

Leimar, O., Hammerstein, P., 2001a. Evolution of cooperation through indirect reciprocity. Proc. R. Soc. Lond. B Biol. Sci. 268 (1468), 745–753.

Leimar, O., Hammerstein, P., 2001b. Evolution of cooperation through indirect reciprocity. Proc. R. Soc. Lond. B Biol. Sci. 268 (1468), 745–753.

Lihoreau, M., Chittka, L., Raine, N.E., 2010. Travel optimization by foraging bumblebees through readjustments of traplines after discovery of new feeding locations. Am. Nat. 176 (6), 744–757.

Lupatsch, I., Kissil, G.W., Sklan, D., 2001. Optimization of feeding regimes for European sea bass Dicentrarchus labrax: a factorial approach. Aquaculture 202 (3–4), 289–302.

Majolo, B., Schino, G., Aureli, F., 2012. The relative prevalence of direct, indirect and generalized reciprocity in macaque grooming exchanges. Anim Behav 83 (3), 763–771.

Markvoort, A.J., Sinai, S., Nowak, M.A., 2014. Computer simulations of cellular group selection reveal mechanism for sustaining cooperation. J. Theor. Biol. 357, 123–133.

Marriott, S., 2002. Red in tooth and claw? Images of nature in modern picture books. Child. Lit. Educ. 33 (3), 175–183.

Mathiron, A.G., Pottier, P., Goubault, M., 2019. Keep calm, we know each other: kin recognition affects aggressiveness and conflict resolution in a solitary parasitoid. Anim Behav 151, 103–111.

Midgley, M., 1983. Selfish genes and social Darwinism. Philosophy 58 (225), 365–377.

Molesti, S., Majolo, B., 2017. Evidence of direct reciprocity, but not of indirect and generalized reciprocity, in the grooming exchanges of wild Barbary macaques (Macaca sylvanus). Am. J. Primatol. 79 (9), e22679.

Monod, J., 1974. On chance and necessity. Studies in the Philosophy of Biology. Palgrave, London, pp. 357–375.

Morin, P.A., Moore, J.J., Chakraborty, R., Jin, L., Goodall, J., Woodruff, D.S., 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265 (5176), 1193–1201.

Muir, W.M., 1996. Group selection for adaptation to multiple-hen cages: selection program and direct responses. Poult. Sci. 75 (4), 447–458.

Murphy, G.P., Dudley, S.A., 2009. Kin recognition: competition and cooperation in Impatiens (Balsaminaceae). Am. J. Bot. 96 (11), 1990–1996.

Nee, S., 1989. Does Hamilton's rule describe the evolution of reciprocal altruism? J. Theor. Biol. 141 (1), 81–91.

Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314 (5805), 1560–1563.

Nowak, M.A., Allen, B., 2015. Inclusive fitness theorizing invokes phenomena that are not relevant for the evolution of eusociality. PLoS Biol. 13 (4).

Nowak, M.A., McAvoy, A., Allen, B., Wilson, E.O., 2017. The general form of Hamilton's rule makes no predictions and cannot be tested empirically. Proceedings of the National Academy of Sciences 114 (22), 5665–5670.

Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous populations. Nature 355 (6357), 250–253.

Nowak, M.A., Sigmund, K., 1998. The dynamics of indirect reciprocity. J. Theor. Biol. 194 (4), 561–574.

Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437 (7063), 1291–1298.

Nowak, M.A., Sigmund, K., Leibowitz, M.L., 2000. Cooperation versus competition. Financial Anal. J. 56 (4), 13–22.

Nowak, M.A., Tarnita, C.E., Wilson, E.O., 2010. The evolution of eusociality. Nature 466 (7310), 1057–1062.

Nowak, M., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature 364 (6432), 56–58.

Numbers, R.L., Numbers, K.S., 2003. Religion Red in Tooth and Claw.

O'Malley, M.A., 2015. Endosymbiosis and its implications for evolutionary theory. Proc. Nat. Acad. Sci. 112 (33), 10270–10277.

Ohtsuki, H., Nowak, M.A., 2007. Direct reciprocity on graphs. J. Theor. Biol. 247 (3), 462–470.

Oksanen, L., Lundberg, P., 1995. Optimization of reproductive effort and foraging time in mammals: the influence of resource level and predation risk. Evol. Ecol. 9 (1), 45–56.

Olendorf, R., Getty, T., Scribner, K., 2004. Cooperative nest defence in red–winged blackbirds: reciprocal altruism, kinship or by–product mutualism? Proc. R. Soc. Lond. B Biol. Sci. 271 (1535), 177–182.

Queller, D.C., Strassmann, J.E., 1998. Kin selection and social insects. Bioscience 48 (3), 165–175.

Rachlin, H., 2019. Group selection in behavioral evolution. Behav. Processes 161, 65–72.

Rand, D.G., Nowak, M.A., 2013. Human cooperation. Trends Cogn. Sci. (Regul. Ed.) 17 (8), 413–425.

Robinson, D., Fitter, A., 1999. The magnitude and control of carbon transfer between plants linked by a common mycorrhizal network. J. Exp. Bot. 50 (330), 9–13.

Rosenhouse, J., 2002. PROBABILITY, OPTIMIZATION THEORY, AND EVOLUTION1. Evolution, 56, pp. 1721–1722.

Rothstein, S.I., Pierotti, R., 1988. Distinctions among reciprocal altruism, kin selection, and cooperation and a model for the initial evolution of beneficent behavior. Ethol Sociobiol 9 (2–4), 189–209.

Russell, A.F., Hatchwell, B.J., 2001. Experimental evidence for kin-biased helping in a cooperatively breeding vertebrate. Proc. R. Soc. Lond. B Biol. Sci. 268, 2169–2174, 1481.

Schino, G., Aureli, F., 2010. Primate reciprocity and its cognitive requirements. Evol. Anthropol. Issues News Rev. 19 (4), 130–135.

Schmitz, O.J., 1991. Thermal constraints and optimization of winter feeding and habitat choice in white-tailed deer. Ecography 14 (2), 104–111.

Schweinfurth, M.K., Call, J., 2019. Revisiting the possibility of reciprocal help in non-human primates. Neurosci. Biobehav. Rev. 104, 73–86.

Schweinfurth, M.K., Taborsky, M., 2018. Relatedness decreases and reciprocity increases cooperation in Norway rats. Proc. Biol. Sci. 285 (1874), 20180035.

Seyfarth, R.M., Cheney, D.L., 1984. Grooming, alliances and reciprocal altruism in vervet monkeys. Nature 308 (5959), 541–543.

Silk, J.B., 2002. Kin selection in primate groups. Int. J. Primatol. 23 (4), 849–875.

Silk, J.B., 2006. Practicing Hamilton's rule: kin selection in primate groups. Cooperation in Primates and Humans. Springer, Berlin, Heidelberg, pp. 25–46.

Smith, J.M., 1978. Optimization theory in evolution. Annu Rev Ecol Syst 9 (1), 31–56.

Stiver, K.A., Dierkes, P., Taborsky, M., Lisle Gibbs, H., Balshine, S., 2005. Relatedness and helping in fish: examining the theoretical predictions. Proc. Biol. Sci. 272 (1572), 1593–1599.

Subiaul, F., Vonk, J., Okamoto-Barth, S., Barth, J., 2008. Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. Anim. Cogn. 11 (4), 611–623.

Sutton, R.S., Barto, A.G., 1998. Reinforcement learning: An introduction. The MIT Press.

Swenson, W., Wilson, D.S., Elias, R., 2000. Artificial ecosystem selection. Proc Nat. Acad. Sci. 97 (16), 9110–9114.

Szolnoki, A., Perc, M., 2015. Conformity enhances network reciprocity in evolutionary social dilemmas. J. R. Soc. Interface 12 (103), 20141299.

Szolnoki, A., Perc, M., Szabó, G., Stark, H.U., 2009. Impact of aging on the evolution of cooperation in the spatial prisoner's dilemma game. Phys. Rev. E 80 (2), 021901.

Thébault, E., Fontaine, C., 2010. Stability of ecological communities and the architecture of mutualistic and trophic networks. Science 329 (5993), 853–856.

Traulsen, A., Nowak, M.A., 2006. Evolution of cooperation by multilevel selection. Proc. Nat. Acad. Sci. 103 (29), 10952–10955.

Trivers, R.L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46 (1), 35–57.

Truskanov, N., Emery, Y., Bshary, R., 2020. Juvenile cleaner fish can socially learn the consequences of cheating. Nat Commun 11 (1), 1–9.

Vadas, R.S., Burrows, M.T., Hughes, R.N., 1994. Foraging strategies of dogwhelks, Nucella lapillus (L.): interacting effects of age, diet and chemical cues to the threat of predation. Oecologia 100 (4), 439–450.

Walder, F., Niemann, H., Natarajan, M., Lehmann, M.F., Boller, T., Wiemken, A., 2012. Mycorrhizal networks: common goods of plants shared under unequal terms of trade. Plant Physiol. 159 (2), 789–797.

Wang, Z., Szolnoki, A., Perc, M., 2013. Interdependent network reciprocity in evolutionary games. Sci Rep 3, 1183.

West, S.A., Griffin, A.S., Gardner, A., 2007. Evolutionary explanations for cooperation. Curr. Biol. 17 (16). R661-R672.

West, S.A., Griffin, A.S., Gardner, A., Diggle, S.P., 2006. Social evolution theory for microorganisms. Nat. Rev. Microbiol. 4 (8), 597–607.

Wilkinson, G., 1988. Reciprocal altruism in bats and other mammals. Ethol. Sociobiol. 9 (2–4), 85–100.

Wilson, D.S., Wilson, E.O., 2008. Evolution "for the Good of the Group. Am. Sci. 96 (5), 380–389.

Wilson, E.O., Hölldobler, B., 2005. Eusociality: origin and consequences. Proc. Nat. Acad. Sci. 102 (38), 13367–13371.

Wingreen, N.S., Levin, S.A., 2006. Cooperation among microorganisms. PLoS Biol. 4 (9).

Yamamoto, S., Tanaka, M., 2009. Do chimpanzees (Pan troglodytes) spontaneously take turns in a reciprocal cooperation task? J Comp Psychol 123 (3), 242.

Zahavi, A., 1995. Altruism as a handicap: the limitations of kin selection and reciprocity. J. Avian Biol. 26 (1), 1–3.