The puzzle(s) of cooperation from an evolutionary perspective

Abstract:

What is cooperation? How does it work? How and when does it emerge? Why? For more than a century, researchers from different academic fields have addressed these fascinating questions and tried to solve the evolutionary problem of cooperation, each of them presenting new elements of resolution. This essay aims to, first, articulate the vast literature on cooperation from a solid evolutionary theoretical perspective. A critical synthesis is elaborated around the interrelated concepts of kin selection, direct and indirect reciprocity, network selection, group selection, the concept of punishment and the ontogeny of human cooperation and its components. In a second phase, this structured knowledge is used as a basis for reflecting on current theoretical limits and issues, where biological causes seem to be intertwined with social and cultural mechanisms. Inter-individual, inter-group and inter-cultural variations, norms, fairness, trust, computational systems, forgiveness and beliefs are considered in relation to theoretical edges of the cooperative puzzle. More integrative, naturalistic and observational approaches, in a collaborative and interdisciplinary environment, illustrated for instance by the conceptual modeling of trust used by artificial intelligences, are expected to give rise to new perspectives, at the frontiers of current knowledge.

Keywords: Altruism, Network selection, Punishment, Reciprocity, Reputation.

What is cooperation? How does it work? How and when does it emerge? Why? For more than a century now, researchers of various academic fields have addressed these exciting questions and tried to solve the evolutionary puzzle of cooperation, each one of them presenting new pieces of solution. Nowadays, how far are we from solving this puzzle?

When Darwin (1859) brought forward the theory of natural selection, it gave rise to a fundamental principle that seemed to explain the logic of the whole living world, with the exception of cooperation and other prosocial behaviors. The theory of natural selection implied that any genotypic or phenotypic modification benefiting its holder in its environment may spread in future generations and participate in the species evolution, as the holders can live longer and reproduce more. However, at the individual level, the theory of natural selection could not yet explain such mysteries as the evolution of cooperative behaviors for which an individual incur a cost in order to benefit another or more individuals (Clutton-Brock, 2009). Indeed, on the one hand, following the natural selection principle, if there are no ultimate benefits, the cooperative self-sacrificing individuals would be disadvantaged in their survival, live shorter and reproduce less than others, leading to the disappearance of cooperation in the living world. On the other hand, cooperation among living beings, from bacteria to whales, does not seem on its way to extinction. For humans for instance, cooperation is considered especially widespread, central to its ecological success (Rand, Arbesman & Christakis, 2011) and a base of our societies (Fu, Hauert, Nowak & Wang, 2008). Therefore, either the theory of natural selection missed an element to be complete at this time, or cooperation has benefits for cooperators that we couldn't yet perceive. Thus the evolutionary puzzle of cooperation emerged.

I argue that the articulation of the literature on cooperation from an evolutionary perspective already yields a consistent, yet incomplete, theoretical framework for the emergence of cooperation. More precisely, the following development of the literature proposes a critical synthesis of the mechanisms related to human cooperation. Although human cooperation is only a restricted view of the cooperation happening among living organisms, it has a very rich literature allowing interesting insights and suggestions. This human-based theoretical framework of cooperation does share some mechanisms with other organisms and is expected to inspire or enrich the theoretical construction of various organism-based frameworks. The plurality of perspectives will certainly prove

essential in furthering the limits of our understanding of cooperation. In the continuity of the evolutionary perspective of cooperation, I therefore come back to some debated theoretical aspects and introduce some conceptual limits to be discussed. I argue that theories around cooperation, intensely debated inside the fields of life sciences, would benefit from interdisciplinary perspectives and collaboration.

A blood-related cooperation

A first major piece for solving the puzzle of cooperation was brought by Hamilton's (1967) theory on genetic relatedness, bringing forth the concept of kin selection and the importance of shared genes. A concise way to explain kin selection, given by Nonacs (2011), is that it can be evolutionary pertinent for an individual to "help relatives if the benefit provided (b), prorated by the genetic relatedness of the recipient (r), exceeds the cost to self (c). According to Hamilton's rule: Helping is adaptive if br >c". In other words, Hamilton's view applies the principle of natural selection at the level of genes, where individuals are a sum of genes and blood relatives are a portion of these genes. In the light of this concept, we understand that it can be worth sharing food or shelter, some of the most valuable resources we have, to favor the survival and reproduction of family members as it is partly our survival and reproduction too.

The process of kin selection can be dissected in two mechanisms, kin choice and kin fidelity (Sachs, Mueller, Wilcox & Bull, 2004). Kin choice presupposes that individuals can recognize kin with some same phenotypic traits, which indicates the presence of identical genes between them, and adapt their behavioral response in order to favor recognized kin (Nonacs, 2011; Sachs *et al.,* 2004). An extreme kin choice mechanism, referred to as the greenbeard system, bases kin recognition and related behavior on a very specific trait (like a green beard), indicating unique shared alleles, rather than on a more global relatedness, called "Hamiltonian" relatedness (Keller & Ross, 1998; Nonacs, 2011). Nonacs (2011) and Trivers (2006) illustrate this greenbeard effect with side-blotch lizards (*Uta stansburiana*) where males can either be blue-throated or orange-throated, and blue-throated lizards are observed to preferentially establish territories next to each other, independently of whole-genome relatedness, and act against large and aggressive male orange-throated lizards.

The latter mechanism of kin selection, kin fidelity, involves recognition of kin through environmental cues. For instance, some parent birds recognize their nest but not the

eggs themselves. Some male birds can thus take care of the eggs inside the recognized nest, even if the eggs result from cuckoldry with neighboring males, or both bird parents may take care of parasitic eggs laid by an unrelated individual (Sachs *et al.*, 2004).

The kin selection theory is a consensual explanation for cooperation between blood relatives (Clutton-Brock, 2009); however, following this theory, the puzzle of cooperation remained unsolved between non-kin.

Cooperation is a strategy

Triver's (1971), a few years later, inspired by the work of Hamilton, brought another major piece to the evolutionary puzzle of cooperation. He elaborated an evolutionary theory on altruism, and led to the concept of 'direct reciprocity', illustrated by an example of his own:

> One human being saving another, who is not closely related and is about to drown, is an instance of altruism. Assume that the chance of the drowning man dying is one-half if no one leaps in to save him, but that the chance that his potential rescuer will drown if he leaps in to save him is much smaller, say, one in twenty. Assume that the drowning man always drowns when his rescuer does and that he is always saved when the rescuer survives the rescue attempt. Also assume that the energy costs involved in rescuing are trivial compared to the survival probabilities. Were this an isolated event, it is clear that the rescuer should not bother to save the drowning man. But if the drowning man reciprocates at some future time, and if the survival chances are then exactly reversed, it will have been to the benefit of each participant to have risked his life for the other (Trivers, 1971).

Thus, under certain circumstances, reciprocation can change a detrimental action into a mutually beneficial one. However, following the principle of natural selection where living beings compete to survive and reproduce, the best strategy for an individual would still be to "cheat" and to do not reciprocate, in other words, to be saved and to not take the risk to save the cooperator when the time comes, reaping all benefits of cooperation without any costs. Besides, in large scale societies with high population's mobility, individuals may not meet twice and thus their altruistic actions have low probabilities to be reciprocated. Therefore, the concept of direct reciprocity can only explain non-kin cooperation in a context of expected or sufficient repeated interactions.

The more (cooperators), the merrier

Regarding the issue of non-reciprocators, also considered as defectors, "cheaters" or free-riders, it is important to consider different levels of selection (Nowak, 2012). If one cheater were to compete with one cooperator to survive, it is considered that the cooperator will be exploited and only the defector may survive. On the contrary, if two groups were competing to survive and one group included mostly cooperators whereas the other group had many non-reciprocators, it is considered that only the group of cooperators would survive. Thus, group selection, or multi-level selection, would theoretically shape populations in favor of cooperators, as noted by Darwin (1879; cited from Nowak, 2012):

> There can be no doubt that a tribe including many members who …were always ready to give aid to each other and to sacrifice themselves for the common good, would be victorious over other tribes; and this would be natural selection.

Moreover, inside each social group, individuals can actively select with whom to interact and strengthen their network links with other cooperators, meanwhile breaking network links with defectors (Fehl, van der Post, & Semmann, 2011; Rand *et al.*, 2011). It creates clusters of mutually beneficial cooperators and consequently exclude the defectors from generous networks. This network selection is based on active linking between individuals and requires local, nonrandom and dynamic social networks with frequent updates in order to sustain cooperation (Rand *et al.*, 2011). For instance, neighbors are a local, nonrandom and dynamic social network with a possibility for frequent interactions. And although neighbors might cooperate with everyone at first, they would then refine their network depending on others' reciprocation and later interact preferentially with a selection of neighbors. This clustering of cooperators appears as a self-organized pattern emerging from direct reciprocity (Fehl *et al.,* 2011) and creates an incentive to cooperate (Rand *et al.*, 2011). Therefore, network selection leading to group selection can help to better understand the evolution of cooperation against non-reciprocators in local conditions.

However, in order to grow from tribes to the actual size of our human societies, it is expected to require additional mechanisms promoting extended prosociality among unrelated individuals in order to sustain group cohesion. Boyd and Richerson (2009) consider a theory of cultural group selection that would operate at the same time as

natural group selection, but at a relatively faster rate. Cultural evolution would be a selective pressure enabling large group cooperation through mutual learning, group competitiveness and favorisation of prosocial behaviors. First, mutual learning is the stepping stone of the theory, and accounts for the emergence and success of a group's culture. Indeed, with mutual learning, cooperative individuals can gather and accumulate knowledge, creating a much more efficient and intelligent system that any single individual could become. Then, because cultural selection has a relatively faster evolution rate than genetic selection, a rapid cultural transmission would occur inside each group. This rapid cultural transmission would lead to divergent knowledge and adaptations among local groups, resulting in an increased local competition and favoring behaviors enhancing groups' competitive abilities. Finally, the last basis of the theory of cultural group selection supports that in such culturally evolved social groups, organisms displaying prosocial behaviors should be favored. Indeed, moral systems ensuring cooperation through rewards and punishments would favor the survival and reproduction of prosocial individuals and even possibly lead to the evolution of social emotions such as empathy or shame. Consequently, groups with more cooperators can better adapt to local environments, generate better living conditions, overcome rival groups, grow in population while sustaining their social structure.

Sharing social knowledge

Furthermore, in the context of social groups, the individual benefits of cooperating can accrue in the presence of bystanders or with the testimony of the recipient. Indeed, as a consequence of cooperating, the altruistic individual will gain a reputation of ‘cooperator’ from observers and hearers, and will be more inclined to receive help from others (Nowak, 2006; see Trivers, 2006). On the contrary, one would avoid being altruistic with a reputed defecting individual in order to avoid exploitation (Nowak & Sigmund, 2005). The use of such social information concerning past behaviors, through gossip or observation, is called indirect reciprocity (Mellis & Semmann, 2010; Nowak, 2012). Reputation can be considered to be built in distinct conceptual ways, such as through image scoring or through standing (see Mellis & Semmann, 2010). A basic conceptualization of image scoring considers that cooperating improves the image of the individual by one unit and defecting regresses the image by one unit, independently of the recipient’s reputation. On the other hand, standing would take the recipient's

reputation as well into consideration, and thus, defecting to a bad-reputed individual would not always negatively impact the reputation (Mellis & Semmann, 2010).

An important aspect of indirect reciprocation is that it allows for a theoretical framework explaining cooperation in large-scale conditions with a well-mixed population and a low probability of reciprocation from the recipient (Fu *et al.*, 2008). Indeed, cooperation is observed in one-shot encounters with strangers, or in charity and donation, even though the cooperator is aware of the 'impossibility' of the recipients to reciprocate the altruistic act (Nowak, 2006). Thus, the spreading of a good reputation and its improvement among our social network would be the benefit of such one-sided cooperative behaviors. The impact on reputation can be channeled through at least three different ways: the recipients' testimony to their social network, the bystanders with their gossip, and the helpers telling about their good deeds or donations to their own social network.

Therefore it now seems possible to understand the evolution of cooperation between kin, local non-kin and even between strangers in spite of a defecting strategy. Nevertheless, at least one important mechanism is still missing in order to understand the maintenance of cooperation in social interactions.

Maintaining cooperation

Indeed, individuals must switch between cooperative and defecting strategies depending on the situation and their personal benefit at stake. First of all, defecting can be an individual strategy despite one's cooperative intention. For instance, individuals may sometimes be incapable of helping back a past cooperator because they lack the necessary resources (*e.g.* energy, a ladder) or skills (*e.g.* swimming skills) at the moment, and thus defect. Besides, an individual would have a limited amount of personal resources to invest in cooperation and may not be able to help all encountered individuals or charities in need. However, individuals may also naturally and strategically defect more often than cooperate in the absence of punishment, even though they are in capacity to help or reciprocate. Indeed, as Axelrod and Hamilton (1981) pointed out, defecting can be an evolutionary stable strategy, even in the case of repeated interactions, and can allow one to reap the greatest benefits. Besides, in situations of fleeting encounters in a wide, unrelated and anonymous population, the regulatory mechanisms of direct reciprocity, reputation and exclusion may have lesser

effects. Consequently, one would fall back to cheating because otherwise, a cooperative individual would probably be cheated on. Although this kind of interaction would have considerably grown with the spread of humanity and deterred a globalized world, cooperation persisted. It is conceptualized that punishment can change the final outcome of cheating, rendering it overall less beneficial than cooperating (Raihani *et al.,* 2012). Therefore punishment, or the threat of it, can be necessary for eliciting and maintaining high levels of cooperation in social interactions (Mellis & Semmann, 2010). A kind of punishment, shunning, was already described as a consequence of either network exclusion of defectors or withheld cooperation towards bad-reputed individuals. This punishment is considered non-costly because it is an absence of action (sharing, helping…) from the punisher that deprives the defector from benefits, while the punisher suffers, on the contrary, no present or future loss. Such self-serving cost-free penalization, leading to an end of interaction, are considered as sanctions rather than punishments (Raihani, Thornton & Bshary, 2012). Punishments, on the other hand, are defined to be costly to both the punisher and the punished individual, and should result in delayed benefit for the punisher (Clutton-Brock & Parker, 1995 in Jensen, 2010). Therefore functional punishment is distinguished from other forms of cost-inflicting actions such as aggression, harassment, dominant display, and other behaviors producing immediate benefits for the actor (Jensen, 2010). Costly punishment is supposed to promote future cooperation (Raihani *et al.*, 2012), appears more effective when the participation in the cooperative task is not compulsory (Hauert *et al.*, 2007), and was more described in group interactions rather than in one-shot encounters or repeated interactions involving only two individuals (Henrich, 2006). In the context of group interactions, costly punishment can be considered as an ‘altruistic punishment’ when only one or a few members bear the cost of punishing defectors while it benefits the whole group in sustaining high levels of cooperation (Jensen, 2010).

Becoming cooperative

Alongside the study of the evolution of cooperation, evolutionary theories also benefited from research on proximal causes investigating the development of cooperative abilities (see Warneken, 2018). Experimental studies on adult subjects failed to clarify the ontogeny of cooperation and consequently resulted in a focus on children subjects instead, investigating the onset of cooperation (Olson & Spelke,

2008). Bases of cooperation were found in the cognitive ability of joint attention and in prosocial behaviors such as helping or sharing (Slocombe & Seed, 2019). The first occurrences during early child development were considered around eight months of age in the form of helping behaviors (Leimgruber, Shaw, Santos & Olson, 2012). Around 14 months of age, experimental paradigms elaborated by Warneken and Tomasello (2007) demonstrated selectivity in helping behaviors between situations in which help was requested and situations in which it was not. For sharing behaviors, the onset age was considered at 18 months of age (Warneken, 2018). At 20 months, children accepted to incur a cost for helping (Warneken, Hare, Melis, Hanus & Tomasello, 2007) and around two years-old they generally could infer from situations and behaviors when help was needed (Warneken, 2018). This is also the age around which cooperation with peers emerged, based on imitation (Brownell, Ramani & Zerwas, 2006). Three year-old children increased their coordination with peers (Brownell *et al.*, 2006) and developed a theory of mind (Endedijk, Cillessen, Cox, Bekkering & Hunnius, 2015; Etel & Slaughter, 2019). Then, the age of five was considered the onset age for reputation-based effects on behaviors, furnishing the bases of indirect reciprocity (Warneken, 2018). It was illustrated by the studies of Leimgruber *et al.* (2012) and Fujii, Takagishi, Koizumi and Okada (2015) in which the presence of either an adult observer or a peer, recipient or observer, influenced children's decisions in a prosocial way, increasing generosity or cooperativity. However, the effect of indirect audience only had an influence around nine years-old when children had developed a higher order theory of mind (Takagishi *et al.,* 2015; Warneken, 2018). Concerning punitive and retaliatory behaviors, they are expressed after early prosocial bases. They develop at age three, with the emergence of the concept of fairness and of an inequity aversion (Warneken, 2018). The evolution of costly altruistic punishment arose later in child development, with strong in-group bias around age six, and in a more impartial way at age eight (Jordan, McAuliffe & Warneken, 2014).

Current perspectives

Although children were found steadily more cooperative along their development between three to 14 years-old, House *et al.* (2013) showed in their multicultural study including non-urban societies that it only concerned prosocial behaviors with no personal costs and that on the contrary, the amount of personal sacrifice for prosocial

behaviors varied among societies and cultures. Cooperativeness of individuals appeared to be related to cultural norms, or social norms, learned during childhood from the age of three and only put into actions around age six, with evidence suggesting a causal impact of these norms on the amount of personal investment in cooperation, but remaining to be clearly demonstrated (Baum, Paciotti, Richerson, Lubell & McElreath, 2012; Fehr & Schurtenberger, 2018; House *et al.*, 2013). In the same line, a study from Gachter, Herrmann & Thoni (2008) analyzed the results from an economic cooperative game (repeated public good experiment without and with punishment) played by 1120 undergraduate students across 16 locations distributed into 6 urban (or large-scale) cultures. They found that cooperation varied between individuals, groups and their defined cultures, with a particular increase in the importance of culture in the condition with punishment compared to the condition without.

Norms would have a predominant role in elaborating standards for judging deviation from fairness, leading to punishment (Henrich, 2006; Trivers, 2006). For example, punishing a defector who did not have the resources to reciprocate could be judged unfair by observers, whereas punishing a defector who could have helped but cheated could be judged fair. However, observers would often lack information on resource availability or on previous interactions and would not be able to directly judge the fairness of an interaction. Observers are then suspected to use indirect parameters to judge interactions from unknown individuals, possibly linked to parameters related to the evaluation of trust. These parameters are expected to be phenotypic traits (*e.g.* behavior, appearance, voice,...) representative of a social value for the observer: for instance, if a man is seen in an aggressive interaction, he may be considered with a negative social value for some observers and if these observers had to interact later with the man, it could biased them into suspecting the aggressive individual to cheat in unclear interactions (see Bacharach & Gambetta, 2001 for the use of signs in trust). Interestingly, in the past two decades, the field of artificial intelligence has developed a research area entirely focused on computational mechanisms for trust and reputation in virtual societies. In these environments, isolated agents gather into networks and share information on other agents for avoiding exploitations and thus optimize exchanges through the construction of trust and reputation, based on different conceptual models including beliefs or expectations (Ghasempouri & Ladani, 2019; Sabater & Sierra, 2005). Besides, the ever-growing exploitation of virtual environments by humans appears to rely on artificial intelligence's cooperation and may provide a practical

modelization for the evolution of cooperation. Studying the issues faced by exploitative agents and their adaptive strategies is assuredly of great interest for evolutionary biology as it explores cooperation, reputation and trust from a practical standpoint and mirrors an evolution of cooperation while addressing emerging issues and strengthening their models in our ever-growing use of virtual environments.

Regarding personal investment in cooperation, a more realistic and practical approach would not consider human behaviors with a binary choice between cooperating and defecting but rather as a continuum of potential investment. In other words, individuals invest more or less of their resources for the benefit of others and defectors are not only those investing nothing but also those who invest not enough. Therefore, a normative way defining cheating would be an investment inferior to the population mean (Raihani *et al.,* 2012). Yet, considering how cooperative an action is, related to a reference, is not enough to yield pertinent information and a value must be associated with the behavior in order to respond adaptively. In other words, if an individual becomes aware that an observed behavior is below a threshold, it appears necessary to also know that it is negative in order to act against it. The concept of fairness is likely to be at work in formulating a judgment on cooperative behavior, so that an individual investment lower to the populational reference would be judged unfair, and thus negative. However, what is fair and what is enough cooperative investment is expected to vary across individuals. As pointed out in Fehr and Schmidt's (1999) paper on fairness and cooperation, the population mean may differ between individuals because population-based assessments rely on subjective biases such as which individuals are considered and what relative importance some of them can have. For instance, Fehr and Schmidt (1999) suspected that influential individuals could be used as natural references and that some population's means only considered a sub-population. Therefore, although the concept of fairness is largely cultural and can lead to inter-cultural differences in cooperation (Henrich, 2006), intra-cultural differences are also expected due to the subjective assessment of a population's mean cooperation. As always, individual differences must be considered with tolerance in order to avoid destructive conflicts. Indeed, if individual A cooperates, according to A's threshold of fair investment, but is punished by B for having not been cooperative enough, according to B's higher threshold of fair investment, how unfair would it be for A ! As a consequence, individual A may want to punish B's unfair punishing behavior, which may give rise to an endless loop of reciprocal spite if B retaliates. Thus, besides the direct cost of punishment for the

punisher, an indirect cost of revenge can arise, and suggest that punishing behaviors occur in case of large difference with one's personal cooperative threshold. However, revenge and retaliation could be unrelated to cooperation itself and may be related to dominance, social status, or pride.

Yet, reciprocal spite loops could theoretically emerge in extreme situations and never cease unless regulatory mechanisms exist. Regarding this issue, Trivers (2006) suggested that the ability to forgive may act against the strive for revenge and retribution, and would be "potentially saving enormous amounts of energy, both outward-directed and inner-consumed" (Trivers, 2006). Forgiveness may play an important role in the evolution of social behaviors in general, considering the existence of mistakes in everyday life. Theoretically, unconditional or spontaneous forgiveness would further allow exploitation and would consequently not be adaptive, contrary to the evolution of conditional mechanisms promoting forgiveness, possibly requesting clues of positive intentions from the other. Seeking forgiveness in a system of reciprocal spite might be costly, even by passively enduring spite without reciprocating. Nevertheless, further spite also induce further costs considering that the investment of resources in the reciprocal spite would not be available for cooperative investment and may lead to increased defection in their cooperative network. The individual in reciprocal spite may thus receive sanctions and punishment from individuals of a concerned social group as well as a negative third-party effect through bad reputation.

Nonetheless, although the concept of reputation is pertinent regarding the well-documented 'observer effect' on behaviors, the mechanisms for reputation building are unclear and its concepts remain debated. Consequently, it would be difficult to anticipate what kind of interactions would actually be involved in reputation building. Regarding punishing behaviors, if we consider that reputation is an estimation of how much an individual could contribute for oneself, then punishment, as well as donations, could be 'advertising' for individual inclination in cooperative behaviors. Nevertheless, behaviors such as altruistic punishment in anonymous conditions would not be a candidate for cooperative advertising and are still unclear regarding individual benefits. Interestingly, Trivers (2006) reported a potential impact of God as an ever-present observer and Purzycki *et al.* (2016) showed in their multicultural study that beliefs in moralistic, punitive and knowing gods increased impartial behavior, and consequently favored prosociality. Besides religious belief, we could further extend it to a more general conceptual thinking where the representation of our value system

through a symbolic individual (*e.g.* a god, a parent, a mentor, a fictional character,…) can act as an ever-present observer to whom we seek an approval of our behaviors. Despite this altruistic incentive, altruistic behaviors such as anonymous donation or anonymous third-party punishment could still be related to reputation's benefits through the actor's divulgence of actions to a social group (*e.g.* "I gave blood yesterday" or "I couldn't stand this unfair situation this morning"). Actor's testimony of good deeds would improve the actor's reputation in its social network and yield direct or delayed benefits, provided that the testimony is trusted. If so, it is expected that individuals with a rather high reputation of cooperation would be believed of an altruistic deed and would actually do it, whereas individuals without a high reputation might not be trusted by hearers and would not benefit from their altruistic act. Besides cooperative reputation, mechanisms of trust may largely regulate the belief and the benefits from testimonies.

Concluding comments

Overall, the puzzle of cooperation from an evolutionary perspective has made great progress since its formulation more than a century ago and the main picture clearly appears regarding the evolution of cooperation. The cooperation between kin can be understood as a benefit to an extended self at the genetic level, whereas cooperation between unrelated individuals can be explained at the organismic level by benefits from repeated interactions, reputation, third-party observation and punishment. The evolution of cooperation is explained at both the individual level through network selection and at the group level through adaptive mutual learning and prosocial culture. Besides gaining insight in the evolution of cooperation at the population level, its ontogeny has been investigated as well, detailing a rich and coherent individual development. Interesting and fine details of the theoretical framework are still emerging and the field keeps on being stimulated by necessary debates while researchers apply to test and investigate its practical implications, through experimental games and, hopefully, through an increased number of realistic settings. Nevertheless, intense debates and conceptual discussions remain at the edge of the puzzle. Among others, the mechanisms of reputation are still theoretically unclear, regarding its construction, regulation, the involved cognitive process or its intimate interaction with the concept of trust. The existence of reputational mechanisms has laid a theoretical foundation for the explanation of

altruism. However, the concept of altruism has been and remains a challenging consideration, especially in the case of anonymous third-party punishment where individuals incur a cost for others, without seemingly reaping any direct or indirect benefits. Despite the suggestion of cultural mechanisms such as religious beliefs or internalized social constructions, the view remains at the individual level, isolated from its sociocultural background, and possibly lacks a more integrative perspective. The consideration of culture, environmental context and individuality are late to be included in the theoretical framework, possibly due to research effort focused on idealist universal rules, but are expected to be of great benefit for crossing the edges of this puzzle. However, such considerations imply more naturalistic and observational experimental designs in order to reach a sociocultural pertinence even in non-urban small-scale societies. Besides, natural settings and experiments will allow the integration of the individual as a part of its environment, enabling a new level of analysis. In this view, studies investigating behaviors during a cooperative task rather than outcomes of an economic game are expected in order to understand the interactive and constructive nature of cooperation. Furthermore, interdisciplinary communication and collaboration appears essential and rich in novel perspectives, modelisation and understandings. The field of cooperation is sure to converse with areas such as artificial intelligence, anthropology, ethnology, sociology, management, advertising, regarding the concepts of trust, reputation, norms, group dynamics and network selection for instance.

Finally, a last major step, although discreet, would be to agree on the edges of this puzzle and disentangle the various fields of research included into the broad concept of cooperation, such as collaboration, help, compliance and altruism among others, surely involving their own set of mechanisms and functions. Defining the subtle differences in the field of cooperation from an evolutionary perspective shall, rather than setting fences between areas, build bridges and channels between sub-puzzle as well as between disciplines, thus improving scientific communication in order to allow debates to move forward. Definitions are of primary importance and it is always surprising, and yet seems ineluctable, that clearly defining the research area shall come so late; although one cannot properly name a matter broadly unknown.

References

Bacharach, M., & Gambetta, D. (2001). Trust as type detection. In *Trust and deception in virtual societies* (pp. 1-26). Springer, Dordrecht. doi:10.1007/978-94-017-3614-5_1

Baum, W. M., Paciotti, B., Richerson, P., Lubell, M., & McElreath, R. (2012). Cooperation due to cultural norms, not individual reputation. *Behavioural processes*, *91*(1), 90-93. https://doi.org/10.1016/j.beproc.2012.06.001

Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a Social Partner With Peers: Cooperation and Social Understanding in One- and Two-Year-Olds. *Child Development, 77*(4), 803–821. https://doi.org/10.1111/j.1467-8624.2006.t01-1-.x-i1

Darwin, C. 1859 On the origin of species by means of natural selection. *London, UK: John Murray*.

Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature*, *462*(7269), 51–57. doi:10.1038/nature08366

Clutton-Brock, T. H. & Parker, G. A. (1995) Punishment in animal societies. *Nature 373,* 209-216. doi:10.1038/373209a0

Endedijk, H. M., Cillessen, A. H. N., Cox, R. F. A., Bekkering, H., & Hunnius, S. (2015). The Role of Child Characteristics and Peer Experiences in the Development of Peer Cooperation. *Social Development, 24*(3), 521–540. https://doi.org/10.1111/sode.12106

Etel, E., & Slaughter, V. (2019). Theory of mind and peer cooperation in two play contexts. *Journal of Applied Developmental Psychology, 60*, 87–95. https://doi.org/10.1016/j.appdev.2018.11.004

Fehl, K., van der Post, D. J., & Semmann, D. (2011). Co‐evolution of behaviour and social network structure promotes human cooperation. *Ecology letters*, *14*(6), 546-551. https://doi.org/10.1111/j.1461-0248.2011.01615.x

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour, 2*(7), 458–468. https://doi.org/10.1038/s41562-018-0385-5

Fu, F., Hauert, C., Nowak, M. A., & Wang, L. (2008). Reputation-based partner choice promotes cooperation in social networks. *Physical Review E*, *78*(2),026117. doi:10.1103/PhysRevE.78.026117

Fujii, T., Takagishi, H., Koizumi, M., & Okada, H. (2015). The Effect of Direct and Indirect Monitoring on Generosity Among Preschoolers. *Scientific Reports*, 5(1). https://doi.org/10.1038/srep09025

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1553), 2651-2661. https://doi.org/10.1098/rstb.2010.0135*

Ghasempouri, S. A., & Ladani, B. T. (2019). Modeling trust and reputation systems in hostile environments. *Future Generation Computer Systems*, *99*, 571-592. doi:10.1016/j.future.2019.05.017

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology, 7(1),* 17–52. https://doi.org/10.1016/0022-5193(64)90039-6

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., ... & Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767-1770. doi:10.1126/science.1127333

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B. S., McElreath, R., Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences, 110*(36), 14586–14591. https://doi.org/10.1073/pnas.1221217110

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2635-2650. doi:10.1098/rstb.2010.0146

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences, 111*(35), 12710–12715. https://doi.org/10.1073/pnas.1402280111

Keller, L., & Ross, K. G. (1998). Selfish genes: a green beard in the red fire ant. *Nature*, *394*(6693), 573-575. doi:10.1038/29064

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young Children Are More Generous When Others Are Aware of Their Actions. *PLoS ONE, 7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Melis, A. P., & Semmann, D. (2010). How is human cooperation different?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2663-2674. doi:10.1098/rstb.2010.0157

Nonacs, P. (2011). Kinship, greenbeards, and runaway social selection in the evolution of social insect cooperation. *Proceedings of the National Academy of Sciences*, *108*(Supplement 2), 10808-10815. doi:10.1073/pnas.1100297108

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science, 314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

Nowak, M. A. (2012). Evolving cooperation. *Journal of Theoretical Biology, 299*(2012), 1–8. https://ped.fas.harvard.edu/files/ped/files/dx.doi.org/10.1016/j.jtbi.2012.01.014

Nowak, M., Sigmund, K. (2005) Evolution of indirect reciprocity. *Nature* 437, 1291–1298. https://doi.org/10.1038/nature04131

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*(1), 222–231. https://doi.org/10.1016/j.cognition.2007.12.003

Raihani, N. J., Thornton, A., & Bshary, R. (2012). Punishment and cooperation in nature. *Trends in ecology & evolution*, *27*(5), 288-295. doi:10.1016/j.tree.2011.12.004

Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, *108*(48), 19193-19198. https://doi.org/10.1073/pnas.1108243108

Sabater, J., & Sierra, C. (2005). Review on computational trust and reputation models. *Artificial intelligence review*, *24*(1), 33-60. doi:10.1007/s10462-004-0041-5

Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. (2004). The evolution of cooperation. *The Quarterly review of biology*, *79*(2), 135-160.

Slocombe, K. E., & Seed, A. M. (2019). Cooperation in children. *Current Biology, 29*(11), R470–R473. https://doi.org/10.1016/j.cub.2019.01.066

Takagishi, H., Fujii, T., Koizumi, M., Schug, J., Nakamura, F., & Kameshima, S. (2015). The development of the effect of peer monitoring on generosity differs among elementary school-age boys and girls. *Frontiers in Psychology, 6.* https://doi.org/10.3389/fpsyg.2015.00895

Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology*, *46*(1), 35-57. doi:10.1086/406755

Trivers, R. (2006). Reciprocal altruism: 30 years later. *Cooperation in primates and humans*, 67-83. doi:10.1007/3-540-28277-7_4

Warneken, F. (2018). How Children Solve the Two Challenges of Cooperation. *Annual Review of Psychology, 69*(1), 205–229. https://doi.org/10.1146/annurev-psych-122216-011813

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous Altruism by Chimpanzees and Young Children. PLoS Biology, 5(7), e184. https://doi.org/10.1371/journal.pbio.0050184

Warneken, F., & Tomasello, M. (2007). Helping and Cooperation at 14 Months of Age. *Infancy, 11*(3), 271–294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x