A mutualistic approach to morality

Nicolas Baumard[1,2], Jean-Baptiste André[3] and Dan Sperber[4,5]

[1] Institute of Cognitive and Evolutionary Anthropology, University of Oxford

[2] Philosophy, Politics and Economics Program, University of Pennsylvania, Philadelphia

[3] Laboratoire Ecologie et Evolution, UMR 7625, CNRS - Ecole Normale Supérieure, Paris

[4] Institut Jean Nicod, ENS, EHESS, CNRS, Paris

[5] Department of Cognitive Science and Department of Philosophy, Central European University, Budapest

Short abstract

What makes humans moral beings? This question can be understood either as a proximate 'how' question or as an ultimate 'why' question. The 'how' question is about the mental and social mechanisms that produce moral judgments and interactions, and has been investigated by psychologists and social scientists. The 'why' question is about the fitness consequences that explain why humans have morality, and has been discussed by evolutionary biologists in the context of the evolution of cooperation. Our goal here is to contribute to a fruitful articulation of such proximate and ultimate explanations of human morality. We do so by developing an approach focusing on recent developments in the study of mutualistic forms of cooperation and on their relevance to fairness-based morality.

Long abstract

What makes humans moral beings? This question can be understood either as a proximate 'how' question or as an ultimate 'why' question. The 'how' question is about the mental and social mechanisms that produce moral judgments and interactions, and has been investigated by psychologists and social scientists. The 'why' question is about the fitness consequences that explain why humans have morality, and has been discussed by evolutionary biologists in the context of the evolution of cooperation. Our goal here is to contribute to a fruitful articulation of such proximate and ultimate explanations of human morality. We develop an approach to morality as an adaptation to an environment in which individuals were in competition to be chosen and recruited in mutually advantageous cooperative interactions. In this environment, the best strategy is to treat others with impartiality and to share the costs and benefits of cooperation equally. Those who offer less than others will be left out of cooperation; conversely, those who offer more will be exploited by their partners. In line with this mutualistic approach, the study of a range of economic games involving property rights, collective actions, mutual help and punishment shows that participants' distributions aim at sharing the costs and benefits of interactions in an impartial way. In particular, the distribution of resources is influenced by effort and talent, and the perception of each participant's rights on the resources to be distributed.

1. Introduction

Humans don't just cooperate. They cooperate in a great variety of quite specific ways and have strong views in each case on how it should be done (with substantial cultural variations). In collective actions aimed at a common goal, there is a *right* way to share the benefits: Those who have contributed more should receive more. When helping others, there is a *right* amount to give. One may have the duty to give a few coins to beggars in the street but one does not owe them half of one's wealth, however helpful it would be to them. When people deserve to be punished, there is a right amount of punishment. Most people in societies with a modern penal system would agree that a year in jail is too much for the theft of an apple and not enough for a murder. People have strong intuitions regarding the right way to share the benefits of activity, the right way to help the needy, and the right way to punish the guilty. Do these intuitions, notwithstanding their individual and cultural variability, have a common logic, and, if so, to what extent is this logic rooted in evolved dispositions?

To describe the logic of morality, many philosophers have noted that when humans follow their moral intuitions, they behave *as if* they had bargained with others in order to reach an agreement about the distribution of the benefits and burdens of cooperation (Gauthier, 1986; Hobbes, 1651; Kant, 1785; Locke, 1689; Rawls, 1971; Scanlon, 1998). Morality, these 'contractualist' philosophers argue, is about maximizing the mutual benefits of interactions. The contract analogy is both insightful and puzzling. On the one hand, it well captures the pattern of moral intuitions, and to that extent well explains why humans cooperate, why the distribution of benefits should be proportionate to each co-operator's contribution, why the punishment should be proportionate to the crime, why the rights should be proportionate to the duties, and so on. On the other hand, it provides a mere *as-if* explanation: it is as if people had passed a contract—but since they didn't, why should it be so?

To evolutionary thinkers, the puzzle of the missing contract is immediately reminiscent of the puzzle of the missing designer in the design of life forms, a puzzle essentially resolved by Darwin's theory of natural selection. Actually, two contractualist philosophers, Rawls and Gauthier, have argued that moral judgments are based on a sense of fairness or impartiality that, they suggested, has been naturally selected. Here we explore this possibility in some detail. How can a sense of fairness evolve?

2. Explaining the evolution of morality

<u>2.1. The mutualistic theory of morality</u>

2.1.1 Cooperation and morality

Hamilton (1964) famously classified forms of social interaction between an 'actor' and a 'recipient' according to whether the consequences they entail for actor and recipient are beneficial or costly (with benefits and costs measured in terms of direct fitness). He called behavior that is beneficial to the actor and costly to the recipient (+/-) *selfishness*, behavior that is costly to the actor and beneficial to the recipient (-/+) *altruism*, and behavior that is costly to the actor and costly to the recipient (-/-) *spite*. Following a number of authors (Clutton-Brock, 2002; Emlen, 1997; Gardner & West, 2004; Krebs & Davies, 1993; Ratnieks, 2006; Tomasello et al., in press), we call behavior that is

beneficial to both the actor and the recipient (+/+) *mutualism*.[1] Cooperation is social behavior that is beneficial to the recipient, and hence cooperation can be altruistic or mutualistic.

Not all cooperative behavior, whether it is mutualistic or altruistic, is *moral* behavior. After all, cooperation is common in and across many living species including plants and bacteria, to which no one is tempted to attribute a moral sense. Among humans, kin altruism and friendship are two cases of cooperative behavior that is not necessarily moral (which is not to deny that being a relative or a friend is often highly moralized). Unlike kin altruism, friendship is mutualistic. In both cases however, the degree of cooperativeness is a function of the degree of closeness—genealogical relatedness in the case of parental instinct (Lieberman, Tooby, & Cosmides, 2007), affective closeness typically linked to the force of common interests in the case of friendship (DeScioli & Kurzban, 2009; Roberts, 2005). In both cases, the parent or the friend is typically disposed to favor the offspring or the close friend at the expense of less closely related relatives or less close friends, and relatives and friends at the expense of third parties.

Behavior based on parental instinct or friendship is aimed at increasing the welfare of specific individuals to the extent that this welfare is directly or indirectly beneficial to the actor. These important forms of cooperation are arguably based on what Tooby, Cosmides et al. (2009; 2010a, b) have described as a Welfare Trade-Off Ratio (WTR). The WTR indexes the value one places on another person's welfare and the extent to which one is disposed, on that basis, to trade off one's own welfare against the welfare of that person. The WTR between two individual is predicted to be a function of the basic interdependence of their respective fitness (see also Rachlin & Jones, 2008 on social discounting). Choices based on WTR considerations typically lead to favoritism and are quite different from choices based on fairness and impartiality. Fairness may lead individuals to give resources to people whose welfare is of no particular interest to them or even to people whose welfare is detrimental to their own. To the extent that morality implies impartiality,[2] parental instinct and friendship are not intrinsically moral.

Forms of cooperation can evolve without morality, but it is hard to imagine how morality could evolve without cooperation. The evolution of morality is appropriately approached within the wider framework of the evolution of cooperation. Much of the recent work on the evolution of human altruistic cooperation has focused on its consequences for morality, suggesting that human morality is first and foremost altruistic (Gintis, Bowles, Boyd, & Fehr, 2003; Haidt, 2007; Sober & Wilson, 1998). Here we focus on the evolution and consequences of mutualistic cooperation. Advances in comparative psychology suggest that, during their history, humans evolved new skills and motivations for collaboration (intuitive psychology, social motivation, linguistic communication) not possessed by other great apes (Tomasello et al., in press). We argue that morality may be seen

---

[1] Note that, from an evolutionary point of view, costs and benefits should be measured over the lifetime. Hence behavior that might seem altruistic when considered in the short run may bring later benefits to the actor and hence be mutualistic.

[2] Of course, there is no generally agreed-upon definition of morality, and it may be argued that morality does not necessarily imply fairness and may include a greater variety of forms of interaction that nevertheless have relevant commonalities (e.g. Haidt, Koller, & Dias, 1993; Shweder, Mahapatra, & Miller, 1987). Here, we use morality in a sense that implies fairness, on the assumption that such a sense picks out a set of phenomena worthy of scientific inquiry, in particular from an evolutionary point of view. Baumard & Sperber (forthcoming) discuss the relation of morality so understood to wider systems of cultural norms.

as a consequence of these cooperative interactions and emerged to guide the distribution of gains resulting from these interactions. (Baumard, 2008). Note that these two approaches are not mutually incompatible. Humans may well have both altruistic and mutualistic moral dispositions. While a great deal of important research has been done in this area in recent decades, we are still far from a definite picture of the evolved dispositions underlying human morality. Our goal here is to contribute to a rich ongoing debate by highlighting the relevance of the mutualistic approach.

2.1.2 The evolution of cooperation by partner choice

Corresponding to the distinction between altruistic and mutualistic cooperation, there are two classes of models of the way in which cooperation may have evolved. Altruistic models describe the evolution of a disposition to engage in cooperative behavior even at a cost to the actor. Mutualistic models describe the evolution of a disposition to engage in cooperation that is mutually beneficial to actor and recipient (see figure 1).
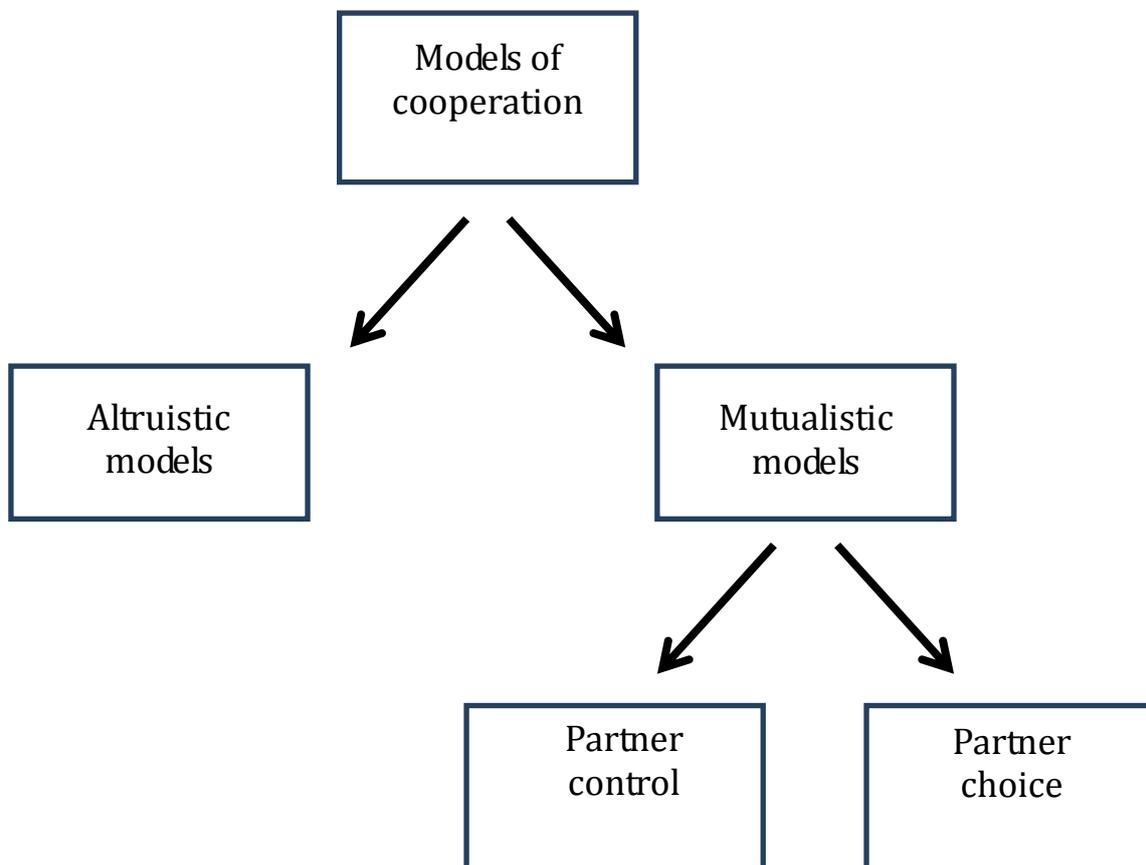

Figure 1: Evolutionary models of cooperation

Mutualistic models are themselves of two main types: those focusing on 'partner control' and those focusing on 'partner choice' (Bshary & Noë, 2003).[3] Earlier mutualistic models were of the

---

[3] There are in principle other possibilities such as by-product mutualism (e.g. group enhancement, pseudo-reciprocity, see Clutton-Brock, 2009) but they are usually not considered in the explanation of human moral behavior.

first type, drawing on the notion of reciprocity as defined in game theory (Raiffa & Luce, 1957; for a review, see Aumann, 1981) and as introduced into evolutionary biology by Trivers (1971).[4] These early models used as their paradigm case iterated Prisoner's Dilemma games (Axelrod, 1984; Axelrod & Hamilton, 1981). Participants in such games who at any time fail to cooperate with their partners can be penalized by them in subsequent trials as in Axelrod's famous 'tit-for-tat' strategy, and this way of controlling one's partner might in principle stabilize cooperation.

In partner control models, partners are given rather than chosen and preventing them from cheating is the central issue. By contrast, in more recently developed partner choice models, individuals can choose their partners and the emphasis is less on preventing cheating than in choosing and being chosen as the right partner (Bull & Rice, 1991; Noë, van Schaik, & Van Hooff, 1991; Roberts, 1998).[5] Consider, as an illustration, the relationship of the cleaner fish *Labroides dimidiatus* with client reef fish. Cleaners may cooperate by removing ectoparasites from clients, or they may cheat by feeding on client mucus. As long as the cleaner eats just ectoparasites, both fish benefit from the interaction. When, on the other hand, a cleaner fish cheats and eats mucus, field observations and laboratory experiments suggest that clients respond by switching partners, fleeing to another cleaner, and thereby creating the conditions for the evolution of cooperative behavior among cleaners (Adam, 2010; Bshary & Grutter, 2005). Reciprocity can thus be shaped either by partner-choice or by partner-control only.

Mutually beneficial cooperation might in principle be stabilized either by partner control or by partner choice (or, obviously, by some combination of both). Partner control and partner choice differ from one another with respect to their response to uncooperativeness, which is generally described as 'defection' or 'cheating'. In partner control models, a cooperator reacts to a cheating partner by cheating as well, thereby either causing the first cheater to return to a cooperative strategy or turning the interaction into an unproductive series of defections. In partner choice models, on the other hand, a co-operator reacts to a partner's cheating by starting a new cooperative relationship with another hopefully more cooperative partner. Whereas in partner control models, individuals only have the choice between cooperating and not cooperating with their current partner, in partner choice model, individuals have the 'outside option' of cooperating with someone else. This difference has, we will see, major implications.[6]

The case of cleaner fish illustrates another important feature of partner choice. In partner choice models, the purpose of switching to another partner is not to inflict a cost on the cheater and thereby punish him. It need not matter to the switcher whether or not the cheater suffers as a consequence. A client fish switching partner is indifferent to the fate of the cleaner it leaves behind. All it wants in switching partner is to benefit from the services of a better cleaner. Still cheating is generally made costly by the loss of opportunities to cooperate at all and this may well have a dissuasive effect and contribute to stabilize cooperation. The choice of new partners is particularly advantageous when it can be based on information about their past behavior. Laboratory

---

[4] Trivers described his own model of mutually beneficial reciprocal interactions as 'reciprocal altruism', but this has been a source of confusion since what is involved is a form of mutualism and hence not of altruism as ordinarily understood.

[5] Note that Trivers shortly discusses this possibility in his foundational article (1971) but does not pursue it.

[6] This difference is similar to Hirschman's influential contrast between 'voice' and 'exit' as the two possible responses available to dissatisfied social or economic actors (Hirschman 1970).

experiments show that reef fish clients gather information about cleaners' behavior and that, in response, cleaners behave more cooperatively in the presence of a potential client (Bshary & Grutter, 2006).

The evolution of cooperation by partner choice can be seen as a special case of 'social selection'. Social selection is a form of natural selection where the selective pressure comes from the social choices of other individuals (Dugatkin, 1995; Nesse, 2007; West-Eberhard, 1979). Sexual selection by female choice is the best-known type of social selection. Female bias for mating with ornamented males select for more elaborate male displays, and the advantages of having sons with extreme displays (and perhaps advantages from getting good genes) select for stronger preferences (Grafen, 1990). Similarly, a socially widespread preference for reliable partners selects for psychological dispositions that foster reliability. When we talk of social selection in the rest of this article, we always refer to the special case of the social selection of dispositions to cooperate.

2.1.3 The importance of partner choice in humans

In humans, many historical and social science studies have demonstrated that partner choice can enforce cooperation without coercion or punishment (McAdams, 1997). European medieval traders (Greif, 1993), Jewish New York jewelers (Bernstein, 1992) and Chinese middlemen in South Asia (Landa, 1981) have been shown, for instance, to exchange highly valuable goods and services without any binding institutions. What deters people from cheating is the risk of not being chosen as partners in future transactions.

In recent years, a range of experiments have confirmed the plausibility of partner choice as a mechanism capable of enforcing human cooperative behavior. They demonstrate that people tend to select the most cooperative individuals, and that those who contribute less than others are gradually left out of cooperative exchanges (Barclay, 2004, 2006; Barclay & Willer, 2007; Chiang, 2010; Coricelli, Fehr, & Fellner, 2004; Ehrhart & Keser, 1999; Hardy & Van Vugt, 2006; Page, Putterman, & Unel, 2005; Rockenbach & Milinski, 2011; K. M. Sheldon, Sheldon, & Osbaldiston, 2000; Sylwester & Roberts, 2010). Further studies show that people are quite able to detect the cooperative tendencies of their partners. They rely on cues such as their partners' apparent intentions (Brosig, 2002), the costs of their actions (Ohtsubo & Watanabe, 2008), or the spontaneity of their behavior (Verplaetse, Vanneste, & Braeckman, 2007). They also actively seek these types of information and are willing to incur costs to get it (Kurzban & DeScioli, 2008; Rockenbach & Milinski, 2011).

A recent experiment well shows how humans have the psychological dispositions necessary for effective partner choice (Pradel, Euler, & Fetchenhauer, 2008). One hundred twenty-two students of six secondary school classes played an anonymous 'dictator game' (see below), which functioned as a measure of cooperation. Afterwards and unannounced, the students had to estimate what their classmates' decisions had been and they did so better than chance. Sociometry revealed that the accuracy of predictions depended on social closeness. Friends (and also classmates who were disliked) were judged more accurately than others. Moreover, the more cooperative participants tended to be friends with one another. There are two prerequisites for the evolution of cooperation through social selection: the predictability of moral behavior and the mutual association of more cooperative individuals. These experimental results show that these prerequisite are

typically satisfied. In a market of cooperative partners, the most cooperative individuals end up interacting with each other and enjoy more common good.

Did human ancestral ecology meet the required conditions for the emergence of social selection? Work on contemporary hunter-gatherers suggests that such is indeed the case. Many studies have shown that hunter-gatherers constantly exchange information about others (Cashdan, 1980; Wiessner, 2005), and that they accurately distinguish good co-operators from bad co-operators (Tooby, Cosmides, & Price, 2006). Field observations also confirm that hunter-gatherers actively choose and change partners. For instance, Woodburn notes that among the Hazda of northern Tanzania, "Units are highly unstable, with individuals constantly joining and breaking away, and it is so easy to move away that one of the parties to the dispute is likely to decide to do so very soon, often without acknowledging that the dispute exists" (Woodburn, 1982, p. 252). Inuit groups display the same fluidity: "Whenever a situation came up in which an individual disliked somebody or a group of people in the band, he often pitched up his tent or build his igloo at the opposite extremity of the camp or moved to another settlement altogether" (Balicki, 1970). Studying the Chenchu, Fürer-Haimendorf (1967) notes that the cost of moving away may be enough to force people to be moral:

> "Spatial mobility and the 'settling of disputes by avoidance' allows a man to escape from social situations made intolerable by his egoistic or aggressive behaviour, but the number of times he can resort to such a way out is strictly limited. There are usually two or three alternative groups he may join, and a man notorious for anti-social behaviour or a difficult temperament may find no group willing to accept him for any length of time. Unlike the member of an advanced society, a Chenchu cannot have casual and superficial relations with a large number of persons, who may be somewhat indifferent to his conduct in situations other than a particular and limited form of interaction. He has either to be admitted into the web of extremely close and multi-sided relations of a small local group or be virtually excluded from any social interaction. Hence the sanctions of public opinion and the resultant approval or disapproval are normally sufficient to coerce individuals into conformity (…)" (p.21)

In a review of the literature on the food exchanges of hunter-gatherers, Gurven (2004) shows that people choose their partners on the basis of their willingness to share (see for instance Aspelin, 1979; Henry, 1951; J. A. Price, 1975). As Kaplan and Gurven (2001) put it, cooperation may emerge from the fact that people in hunter-gatherer societies "vote with [their] feet." (on this point, see also Aktipis, 2004). Overall, anthropological observations strongly suggest that social selection may well have taken place in the ancestral environment.

2.1.4 Outside options constrain the outcome of mutually advantageous interactions

Although mutualistic interactions have evolved because they are beneficial to every individual participating in these interactions, they nonetheless give rise to a conflict of interest regarding the quantitative distribution of payoffs. As Rawls (1971) puts it:

> "Although a society is a cooperative venture for mutual interest, it is typically marked by a conflict as well as by an identity of interest. There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to live

solely by his own efforts. There is a conflict of interests since persons are not indifferent as to how the greater benefits produced by their collaboration are distributed, for in order to pursue their ends they each prefer a larger to a lesser share." (p. 126)

How may such a conflict of interest be resolved among competing partners? There are many ways to share the surplus benefit of a mutually beneficial exchange, and models of "partner control" are of little help here. These models are notoriously under-determined (a symptom of what game theoreticians call the « folk theorem, » e.g. Aumann & Shapley, 1992). This can be easily understood. Almost everything is better than being left without a social interaction at all. Therefore, when the individuals engaged in a social interaction have no outside options, it is generally more advantageous for them to accept the terms of the interaction they are part of than to reject the interaction altogether and be left alone. In particular, even highly biased and unfair interactions may well be evolutionarily stable in this case.

What is more, when individuals have no outside options, the allocation of the benefits of cooperation is likely to be determined by a power-struggle. The fact that an individual has contributed this or that amount to the surplus benefit of the interaction need not have any influence on that power- struggle, nor on the share of the benefit this individual will obtain. In particular, if a dominant individual has the ability to commit to a given course of interaction, then the others will have no better option than to accept it, however unfair it might be (Schelling, 1960). Quite generally, in the absence of outside options, there is no particular reason why an interaction should be governed by moral considerations. There is no intrinsic property of partner control models of cooperation that would help explain the evolution of fairness and impartiality.

On the other hand, fairness and impartiality can evolve when partner-choice rather than partner-control is at work (Baumard, 2008). Using numerical simulations in which individuals can choose with whom they wish to interact, Chiang (2008) has observed the emergence of fairness in an interaction in which partner-control alone would have led to the opposite. André and Baumard (2011) develop a formal understanding of this principle in the simple case of a pairwise interaction. Their demonstration is based on the idea that negotiation over the distribution of benefits in each and every interaction is constrained by the whole range of outside opportunities, determined by the market of potential partners. When social life is made up of a diversity of opportunities in which one can invest time, resources, and energy, one should never consent to enter an interaction in which the marginal benefit of one's investment is lower than the average benefit one could receive elsewhere. In particular, if all the individuals involved into an interaction are 'equal', not in the sense that they have the same negotiation power within the interaction, but in the more important sense that they have the same opportunities outside the interaction, they should all receive the same marginal benefit from each resource unit that they invest in a joint cooperative venture, irrespective of their local negotiating power. Even in interactions in which it might seem that dominant players could get a larger share of the benefits, a symmetric bargaining always occurs at a larger scale, in which each player's potential opportunities are involved.

A biological way of understanding this result is to use the concept of resource allocation. When individuals can freely choose how to allocate their resources across various social opportunities throughout their lives, biased splits disfavoring one side in an interaction are not evolutionarily stable because individuals then refuse to enter into such interactions when they

happen to be on the disfavored side. This can be seen as a simple application of the marginal value theorem to social life (Charnov, 1976). In evolutionary equilibrium, the marginal benefit of a unit of resource allocated to each possible activity (reproduction, foraging, somatic growth, etc.) must be the same. In the social domain this entails, in particular, that the various sides of an interaction must benefit in the same manner from this interaction, otherwise one of them is better off refusing.

This general principle leads to precise predictions regarding the way social interactions should take place. We have just explained that individuals should share their common goods *equally* when they have contributed *equally* to their production. However, in many real life instances, individuals play distinct roles, and participate differently in the production of a common good. In this general case, we suggest that they should be rewarded as a function of the *effort* and *talent* they invest into each interaction. Let us explain why.

First, if a given individual, say A, participates in an interaction in which he needs to invest say three "units of resources", whereas B's role only involves investing one unit, then A should receive a payoff exactly three times greater than B. If A's payoff is less than three time's B, then A would be better off refusing, and playing three times B's role in different interactions (e.g. with other partners). Individuals should always receive a net benefit *proportional* to the amount of resources they have invested in a cooperative interaction so that they benefit *equally* from the interaction (given their investment). This, incidentally, corresponds in moral philosophy to Aristotle's proportionality principle.

Second, individuals endowed with a special talent, who have the ability to produce larger benefits than others, should receive a larger fraction of the common good. In every potential interaction into which a talented individual can potentially enter, she will find herself in an efficient interaction (an interaction in which at least one player is talented, namely herself), whereas less talented individuals may often find themselves in inefficient ventures. In any given interaction, the average outside opportunities of a talented player are thus higher, and hence she should receive a larger fraction of the benefits, otherwise she is better off refusing to take part in the interaction. Again, individuals benefit *equally* from the interaction given the value of the talents and abilities they invested in the interaction.

In conclusion, mutualistic models of cooperation based on partner control only (e.g. Axelrod & Hamilton, 1981) are unable to generate quantitative predictions regarding the way mutually beneficial cooperation should take place. In contrast, mutualistic models accounting explicitly for the unsatisfied individuals' option of changing partners (André & Baumard, 2011) show that cooperative interactions can only take a very specific form that has all the distinctive features of fairness, defined as *impartiality* or *mutual advantage*. Individuals should be rewarded in exact proportion to the effort they invest in each interaction, and as a function of the quality and rarity of their skills, otherwise they are better off interacting with other partners.

2.2 Three challenges for mutualistic approaches

For a long time, evolutionary theories of human cooperation were dominated by mutualistic theories (Axelrod & Hamilton, 1981; Trivers, 1971). In the last two decades, it has been argued that mutualistic approaches face several problems (Boyd & Richerson, 2005; E. Fehr & Henrich, 2003; Gintis et al., 2003). Three problems in particular have been highlighted: 1) humans spontaneously

help others—even when they have not been helped previously, 2) humans cooperate in anonymous contexts—even when their reputation is not at stake, and 3) humans punish others—even at a cost to themselves. In the following section, we show how a mutualistic theory of cooperation can accommodate these apparent problems.

2.2.1 The evolution of an intrinsic motivation to behave morally

The mutualistic approach provides a straightforward explanation of why people should strive to be good partners in cooperation and respect the rights of others: if they failed to do so, they would incur the risk of being left out of future cooperative ventures. On the other hand, the theory of social selection as stated so far says very little about the proximal psychological mechanisms that are involved in helping individuals compete to be selected as partners in cooperation. In particular, the theory does not by itself explain why humans have a moral sense, why they feel guilty when they steal from others and why they feel outraged when others are treated unfairly (Fessler & Haley, 2003).

In principle, people could behave as good partners and do well in the social selection competition not out of any moral sense and without any moral emotion, but by wholly relying only on self-serving motivations. They could take into account others' interests when this affects their chances of being chosen as partners in future cooperation and not otherwise. They could ignore others' interests when their doing so could not be observed or inferred by others. This is the way intelligent sociopaths tend to behave (Cima, Tonnaer, & Hauser, 2010; Hare, 1993; Mealey, 1995). Sociopaths can be very skilled at dealing with others: they may bargain, make concessions, and be generous, but they only do so in order to maximize their own benefit. They never pay a cost without expectation of a greater benefit. Although normal people also do act morally for self-serving motives and take into account the reputational effect of their actions (see for instance Haley & Fessler, 2005; Hoffman, McCabe, & Smith, 1996; Rigdon, Ishii, Watabe, & Kitayama, 2009), the experimental literature has consistently shown that most individuals—in particular in anonymous situations—commonly respect others' interests even when it is not in their own interest to do so.

The challenge therefore is to explain why, when they cooperate, people have not only *selfish* motivations (that may cause them to respect others' interest for instrumental reasons: getting resources and attracting partners) but also *moral* motivations causing them to respect others' interests per se.

To answer the challenge, it is necessary to consider not only the perspective of an individual wanting to be chosen as a partner but also that of an individual or a group deciding with whom to cooperate. This is an important decision that may secure or jeopardize the success of cooperation. Hence, just as there should have been selective pressure to behave so as to be seen as a reliable partner, there should have been selective pressure to develop and invest adequate cognitive resources in recognizing truly reliable partners. Imagine that you have the choice between two possible partners, call them Bob and Ann, both of whom have, as far as you know, been reliable partners in cooperation in the past. Bob respects the interests of others for the reason and to the extent that it is in his interest to do so. Ann respects the interests of others because she values doing so per se. In other words she has moral motivations. As a result, in circumstances where it might be advantageous to your partner to cheat, Ann is less likely to do so than Bob. This, everything else being equal, makes Ann a more reliable and hence a more desirable partner than Bob.

But how can you know whether a person has moral or merely instrumental motivations? Bob, following his own interest, respects the interest of others either when theirs and his coincide, or when his behavior provides others with evidence of his reliability. Otherwise, he acts selfishly and at the expense of others. As long as he never makes any mistake and behaves appropriately when others are informed of his behavior, the character of his motivations may be hard to ascertain. Still a single mistake, e.g. acting on the wrong assumption that there are no witnesses, may cause others to withdraw their trust and be hugely costly. Moreover, our behavior provides a lot of indirect and subtle evidence of our motivations.

Humans are expert mindreaders. They can exploit a variety of cues, taking into account not only outcomes or interactions but also what participants intentionally or unintentionally communicate about their motivations. Tetlock et al. (2000), for instance, asked people to judge a hospital administrator who had to choose either between saving the life of one boy or another boy (a tragic trade-off where no solution is morally satisfactory), or between saving the life of a boy and saving the hospital $1 million (another trade-off, but one where the decision should be obvious from a moral point of view). This experiment manipulated: (a) whether the administrator found the decision easy and made it quickly, or found the decision difficult and took a long time; (b) which option the administrator chose. In the easy trade-off condition, people were most positive towards the administrator who quickly chose to save Johnny whereas they were most punitive towards the administrator who found the decision difficult and eventually chose the hospital (which suggests that they could sacrifice a boy for a sum of money). In the tragic trade-off condition, people were more positive towards the administrator who made the decision slowly rather than quickly, regardless of which boy he chose to save. Thus, lingering over an easy trade-off, even if one ultimately does the right thing, makes one a target of moral outrage. But lingering over a tragic trade-off serves to emphasize the gravity of the issues at stake and the due respect for each individual's right. More generally, many studies suggest that it is difficult to completely control the image one projects; that there are numerous indirect cues to an individual's propensity to cooperate (Ambady & Rosenthal, 1992; Brown, 2003); and that participants are able to predict on the basis of such cues whether or not their partners intend to cooperate (Brosig, 2002; Frank, Gilovich, & Regan, 1993).

Add to this the fact that people rely not only on direct knowledge of possible partners but also on information obtained from others. Humans—unlike other social animals—communicate a lot about each other through informal gossip (Barkow, 1992; Dunbar, 1993) and more formal public praises and blames (McAdams, 1997). As a result an individual stand to benefit of suffer not only from the opinions that others have formed of her on the basis of direct personal experience and observation but also from a reputation that is being built through repeated transmission and elaboration of opinions that may themselves be based not on direct experience but on others' opinions. A single mistake may compromise one's reputation not only with the partner betrayed but with a whole community. There are, of course, costs of missed opportunities in being genuinely moral and not taking advantages of opportunities to cheat. There may be even greater costs in pursuing one's own selfish interest all the time: high cognitive costs involved in calculating risks and opportunities and, more importantly, risks of incurring huge losses just in order to secure relatively minor benefits. The most cost-effective way of securing a good moral reputation may well consist in being a genuinely moral person.

In a mutualistic perspective, the function of moral behavior is to secure a good reputation as a co-operator. The proximal mechanism that has evolved to fulfill this function is, we argue, a genuine moral sense (for a more detailed discussion, see Baumard & Sperber, submitted).This account is in the same spirit as a well-known argument made by Trivers (1971; see also Frank, 1988; Gauthier, 1986)

> "Selection may favor distrusting those who perform altruistic acts without the emotional basis of generosity or guilt because the altruistic tendencies of such individuals may be less reliable in the future. One can imagine, for example, compensating for a misdeed without any emotional basis but with a calculating, self-serving motive. Such an individual should be distrusted because the calculating spirit that leads this subtle cheater now to compensate may in the future lead him to cheat when circumstances seem more advantageous (because of unlikelihood of detection, for example, or because the cheated individual is unlikely to survive)." (Trivers, 1971, p. 51)

While we agree with Trivers that cooperating with genuinely moral motives may be advantageous, we attribute a somewhat different role to moral motivation in cooperation. In classical mutualistic theories, a moral disposition is typically seen as a psychological mechanism selected to motivate individual to give resources to others. In a mutualistic approach based on social selection like the one we are exploring here, we stress that much cooperation is mutually beneficial so that self-serving motives might be enough to motivate individuals to give resources. Individuals have indeed very good incentive to be fair for if they fail to offer equally advantageous deals to others, they will be left for more generous partners.

What we are arguing is that the function securing a good reputation as a cooperator is more efficiently achieved, at the level of psychological mechanisms, by a genuine moral sense where cooperative behavior is seen as intrinsically good rather than by a selfish concern for one's reputation.[7] Moreover, the kind of cooperative behavior that one has to value in order to achieve useful reputational effects is fairly specific. It must be behavior that makes one a good partner in mutualistic ventures. Imagine for instance a utilitarian altruist willing to sacrifice not only his benefits (or even his life) but also those of his partners for the greater good of the larger group: this might be commended by utilitarian philosophers as the epitome of morality, but, even so, it is not the kind of disposition one looks for in potential partners. Partner choice informed by potential partners' reputation selects for a disposition to be fair rather than for a disposition to sacrifice oneself or for virtues such as purity or piety that are orthogonal to one's value as a partner in most cooperative ventures. This is not to deny that these other virtues may also have evolved either biologically or culturally, but, we suggest, the selective pressures that may have favored them are distinct from those that have favored a fairness-based morality.

Distinguishing a more general disposition to cooperate from a more specific moral disposition to cooperate fairly has important evolutionary implications. Some social traits, for instance Machiavellian intelligence, are advantageous to an individual whether or not they are

---

[7] We do not to deny the existence of that a concern for one's reputation that plays an important role not just in moral behavior, but in all forms of behavior where our reputation may be at stake, for example in matters of skills, intelligence, strength, or sex-appeal (see Baumard & Sperber, submitted).

shared in its population. Other social traits, for instance a disposition to emit and respond to alarm calls, are advantageous to an individual only when they are shared in its population. In this respect, a mere disposition to cooperate, and a disposition to do so fairly belong to different categories. An evolved disposition to cooperate is adaptive only when it is shared in a population: a mutant disposed to give resources to others would be at a disadvantage in a population where no one else would have the disposition to reciprocate. On the other hand, in a population of cooperators competing to be chosen as partners, a mutant disposed to cooperate fairly, not just when it is to her short-term advantage but always, might well be overall advantaged, even if no other individual had the same disposition, because this would enhance her chances to be chosen as a partner. This suggests a two-step account of the evolution of morality:

> Step 1: Partner-choice favors individuals who share equally the costs and benefits of cooperative interactions (see section 2.1.4.). At the psychological level, mutually advantageous reciprocity is motivated by selfish reasons and Machiavellian calculus.

> Step 2: Competition among cooperative partners leads to the selection of a disposition to be intrinsically motivated to be fair (see this section). At the psychological level, mutually advantageous reciprocity is motivated by a genuine concern for fairness.

2.2.2 The scope of cooperation in mutualistic approaches

As we mentioned in section 2.1, early mutualistic approaches to cooperation were focused on partner control in strictly reciprocal relationships, real-life examples of which are provided by barter or the direct exchange of gifts or services. Many cooperative acts, however, like giving a ride, holding doors or giving money to the needy cannot be adequately described in terms of reciprocity so understood. One might of course describe such acts as examples of 'generalized reciprocity', a notion that, according to Sahlins who made it popular, "refers to transactions that are putatively altruistic, transactions on the line of assistance given and, if possible and necessary, assistance returned. […] This is not to say that handing over things in such form […] generates no counter-obligation. But the counter is not stipulated by time, quantity, or quality: the expectation of reciprocity is indefinite" (Sahlins, 1965, p. 147). Such forms of cooperation are indeed genuine and important. Their description in terms of 'reciprocity' however can be misleading if it leads one to expect that the evolution of 'generalized reciprocity' could be explained by generalizing standard models of the evolution of strict reciprocity. Free-riding, which is already difficult to control in strict reciprocity situations, is plainly not controlled in these cases; hence the evolution of 'generalized reciprocity' cannot be explained in terms of partner control. We prefer therefore to talk of 'mutualism' rather than of 'generalized reciprocity'

A first step towards clarifying the issue is to characterize mutualistic forms of cooperation directly rather than by loosening the notion of reciprocity. The clearest real-life examples of mutualistic cooperation is provided by mutual insurance companies (such as the "Philadelphia Contributionship for the Insurance of Houses From Loss by Fire" founded by Benjamin Franklin in 1752) where every member contributes to covering the risks of all members (Broten, 2010; G. Emery & Emery, 1999; Gosden, 1961). All members pay a cost but only some members get an actual benefit. Still, paying the cost of an insurance is *mutually beneficial* to all in terms of expected utility. Generally speaking, we have mutualism when behaving cooperatively towards others is advantageous because of the average benefits that are expected to result for the individual from

such behavior. These benefits can be direct, as in the case of reciprocation, or indirect and mediated by the effect of the individual behavior on her chances to be recruited in future cooperative interactions.

When it is practiced informally, mutualistic cooperation so understood does not allow proper bookkeeping and might therefore seem much more open to free-riding than strict reciprocity. With partner choice however, the reverse may well be the case. Informal mutualistic cooperative actions differ greatly in the degree to which they are individually compensated by specific cooperative actions of others, but nearly all contribute to some degree to the individual's moral reputation. A mutualistic social life is dense with information relevant to everyone's reputation. This information is not confined to cases of full defection (e.g. refusal to cooperate) but is present in all intermediate cases between full defection and perfect equity. Each time one holds the door a bit too shortly or hesitates to give a ride, one signals one's reluctance to meet the demands of mutually advantageous cooperation. More generally, the more varied the forms of informal mutual interaction, the richer the evidence which can sustain or on the contrary compromise one's reputation. Hence being uncooperative in a mutualistic community may be quite costly; 'free-riding' may turn out anything but free. In such conditions, it is prudent—a prudence which, we suggest, is built into our evolved moral disposition— to behave routinely in a moral way rather than check each and every time whether one could, without reputational costs, profit by behaving immorally (Frank, 1988; Gauthier, 1986; Trivers, 1971).

Partner choice combined with reputational information makes possible among humans the evolution of mutualistic cooperation well beyond the special case of reciprocal interactions. How far and how deep can mutualistic relationships extend? This is likely to depend to a large extent on social and environmental factors. But *mutatis mutandis*, the logic is likely to be the same as in mutual aid societies: individual should help each others in a way that is mutually advantageous, i.e. they should help each other to the extent that the cost of helping is smaller less than the expected benefit of being helped when in need. Thus, we will hold doors when people are close to them, but not when they are far away; we will offer a ride to a friend who has lost his car keys, but not drive him home every day; we will help the needy, but only up to a point (the poverty line).

The requirement that interaction should be mutually beneficial limits the forms of help that are likely to become standard social practices. If I can help a lot at a relatively low cost, I should. If, on the other hand, I can only help a little and at a high cost, I need not. In other words the duty to help others depends on the costs (c) to the actor and benefits (b) to the recipient. As in standard reciprocity theories, individuals should only help others when on average b>c. Our obligations to help others are thus limited: we ought to help others only insofar as it is mutually advantageous to do so. Mutual help however can go quite far. Consider for instance a squad of soldiers having to cross a mine field. If each follows his own paths, their chances of surviving are quite low. If on the other hand they walk in line one behind another they divide the average risk. But who should help his comrades by walking in front? Mutuality suggests that they should take equal turns.

The best partners are thus those who adjust their help to the circumstances so as to always behave in mutually advantageous way. This means that evolution will select not only a disposition to

cooperate with others in a mutually advantageous manner, but also a disposition to cooperate with others *whenever* it is mutually advantageous to do so.[8] If, say, offering a ride to a friend who has lost his car keys is mutually advantageous (it helps your friend a lot and does not cost you too much), then if you fail to offer a ride, not just I but also others may quickly figure out that you are not an interesting cooperator. If the best partners are those who always behave in mutually advantageous way, this explains why morality is not only *reactive* (compensating others) but also *proactive* (helping each other, see Elster, 2007 for such as distinction). Indeed, in a system of mutuality, individuals really *owe* others many goods and services—they have to help others—for if they failed to fulfill their duties toward others, they would reap the benefits of mutuality without paying its costs (Scanlon, 1998). This 'proactive' aspect of much moral behavior is responsible for the 'illusion' that individuals act as if they had previously agreed on a contract with each other.

2.2.3 Punishment from a mutualistic perspective

Recently, models of altruistic cooperation and experimental evidence have been used to argue that punishment is typically altruistic, often meted out at a high cost to the punisher, and that it evolved as an essential way to enforce cooperation (Boyd, Gintis, Bowles, & Richerson, 2003; Gintis et al., 2003; J. Henrich et al., 2006). In partner choice models, on the other hand, cooperating is made advantageous not so much by the cost punishment in the case of non-cooperation, as by the need to attract potential partners.

There is much empirical evidence that is consistent with the mutualistic approach. Punishment as normally understood (that is behavior that not only imposes a cost but has the function of doing so) is uncommon in societies of foragers (see Marlowe, 2009 for a recent study) and, in these societies, most disputes are resolved by self-segregation (Guala, 2010, see also Baumard 2010a for a review). In most cases, people simply stop wasting their time interacting with immoral individuals. If the wrongdoing is very serious and threatens the safety of the victim, she may retaliate in order to preserve her reputation or deter future aggressions (McCullough, Kurzban, Tabak, Shaver, & Mikulincer, 2010). Such behavior, however, cannot be seen as punishment *per se* since it is aimed only at protecting the victim (as in many non human species, Clutton-Brock & Parker, 1995). Furthermore, although punishment commonly seeks to finely rebalance the interests of the wrongdoer and the victim, retaliation can be totally disproportionate and much worse than the original aggression (Daly and Wilson, 1988). There is, in fact, good evidence that people in small-scale societies distinguish between legitimate (and proportionate) punishment and illegitimate (and disproportionate) retaliation (Fürer-Haimendorf, 1967; Miller, 1990).[9]

Although humans, in a mutualistic framework, have not evolved an instinct to punish, some punishment is to be expected though in three cases: 1) when the victim is also the punisher and has

---

[8] Incidentally, this analysis explains why someone who always reciprocates can still be morally condemned. Indeed, being moral is not about strict reciprocity but about mutual advantage. Let's say that John always pay his debts, keeps note of every penny lend by friends, reciprocate every glass of wine shared in a bar but never help someone he does not know and will not meet again, even when it is almost costless to him. We would not like to have John as our friend. Reciprocating on a strict act basis is evidence of a lack of cooperative dispositions. The true moral behavior is to help others when they need, and they (or others) will help you when you need it.
[9] Punishment proper is much more important in large societies (Black, 2000), but it is carried through specialized institutions that reward people for the job of punishing (via gratification or policing, Ostrom, 1990).

a direct interest in punishing (punishment then coinciding with retaliation or revenge), 2) when the costs of punishing are incurred not by the punishers but by the organization—typically the state— that employs and pays them, and 3) when the cost of punishing is negligible or insignificant (indeed, in this case, refraining from punishing would amount to be an accomplice to the wrongdoing). In these cases, punishment does not serve to deter cheating as in altruistic theories, but to restore fairness and balance between the wrongdoer and the victim.

How costly should the punishment be to the person punished? Basically, from a mutualistic point of view, it should be high enough to re-establish fairness but low enough not to create additional injustice (by harming the wrongdoer disproportionately). The guilty party who has harmed or stolen from others should, if at all possible, compensate his victims and should suffer in proportion to the advantage he had unfairly sought to enjoy. Here, punishment involves both restorative and retributive justice and is the symmetric counterpart of distributive justice and mutual help. Just as people give resources to others when others are entitled to them or need them, people take away resources from those who were not entitled to them and impose a cost on them that is proportionate to the benefit they might have unfairly enjoyed.

2.3. Predictions

Is human cooperation governed by a principle of impartiality? In this section, we spell out the predictions of the mutualistic approach in more detail and examine whether they fit with moral judgments.

2.3.1 Exchange: Proportionality between contributions and distributions

Human collective actions, for instance collective hunting or collective breeding, can be seen as ventures in which partners invest some of their resources (goods and services) to obtain new resources (e.g. food, shelter, protection) that are more valuable to them than the ones they have initially invested. Partners, in other words, offer their contribution in exchange for a share of the benefits. For this, partners need to assess the value of each contribution, and to proportionate the share of the benefits to this value. In section 2.1.4, we outlined the results of an evolutionary model predicting that individuals should share the benefits of social interactions *equally*, when they have *equally* contributed to their production (André & Baumard, 2011). We saw that this logic predicts that participants should be rewarded as a function of the *effort* and *talent* that they invest in each interaction (although a formal proof will require further modeling).

Experimental evidence confirms this prediction, showing a widespread massive preference for meritocratic distributions: the more valuable your input, the more you get (Konow, 2001; Marshall, Swift, Routh, & Burgoyne, 1999). Similarly, field observations of hunter-gatherers have shown that hunters share the benefits of the hunt according to each participant's contribution (Gurven, 2004). Bailey (Bailey, 1991), for instance, reports that in group hunts among the Efe Pygmies, initial game distributions are biased toward members who participated in the hunt, and that portions are allocated according to the specific hunting task. The hunter who shoots the first arrow gets an average of 36% , the owner of the dog who chased the prey gets 21%, and the hunter who shoots the second arrow gets only 9% by weight (see also Alvard & Nolin, 2002 for the distribution of benefits among whale hunters).

Social selection should, moreover, favor considerations of fairness in assessing each partner's contribution. For instance, most people who object to CEOs' or football stars' huge salaries do so not out of simple equalitarianism but because they see these salaries as far above what would be justified by the actual contributions to the common good of those who earn them (for an experimental approach, see Konow, 2003). Such assessments of individual contributions are themselves based, to a large extent, on the assessor's understanding of the workings of the economy and of society. As a result, a similar sense of fairness may lead to quite different moral judgments on actual distribution schemes. Europeans, for instance, tend to be more favorable to redistribution of wealth than Americans. This may be not because they care more about fairness but because they are more likely to think of the poor as being unfairly treated (Alesina & Glaeser, 2004; on the relationships between belief in meritocracy and judgement on redistribution, see also Fong, 2001; on factual beliefs about wealth distribution see Norton & Ariely, 2011). In other words, when Americans and Europeans disagree on what the poor deserve, their disagreement may stem from their understanding of society rather than from the importance they place on fair distribution.

2.3.2 Mutual aid: Proportionality between costs and benefits

In the kind of collective actions we just discussed, the benefits are distributed in proportion to individual contributions. As we already insisted, mutualistic cooperation in not limited to such cases. In mutual aid in particular, contributions are based on the relationship between one's cost in helping to help and one's benefit in being helped. Mutual aid may be favored over strict reciprocity for several reasons: for instance, because risk levels are high and mutual aid provides insurance against them, or because individuals cooperate on a long-term basis where everyone is likely to be in turn able to help or in need of help. Mutual aid is widespread in hunter-gatherer societies (Barnard & Woodburn, 1988; see Gurven, 2004 for a review). Among Ache (Kaplan & Gurven, 2001; Kaplan & Hill, 1985), Maimande (Aspelin, 1979), and Hiwi (Gurven, Hill, Kaplan, Hurtado, & Lyles, 2000), shares are distributed in proportion to the number of consumers within the recipient family. Families with high dependency tend to be net consumers, whereas those with low dependency are net producers among the Batak (Cadelina, 1982). Among the G/wi, the largest shares of game are first given to families with dependent children, then to those without children, and the smallest shares are given to single individuals (Silberbauer, 1981).

Note that mutual aid is quite compatible with the kind of meritocratic distributions observed in collective actions. Among hunter-gatherers, non-meat items and cultigens whose production is highly correlated with effort are often distributed according to merit, while meat items whose production is highly unpredictable are distributed much more equally (Alvard, 2004; Gurven, 2004; Wiessner, 1996). Similarly, it is often possible in hunter-gatherer societies to distinguish the primary distribution based on merit, in which hunters are rewarded as a function of their contribution to the hunt, and the secondary distribution which is based on need, in which the same hunters share their meat with their neighbors thereby obtaining insurance against adversity (Alvard, 2004; Gurven, 2004).

Of course, the help we owe to others varies according to circumstances, and from society to society. Higher levels of mutual aid are typically observed among relatives or close friends because their daily interactions and their long-term association make mutual aid less costly, more advantageous, and more likely to be reciprocated on average (Clark & Jordan, 2002; Clark & Mills,

1979). It also depends on the type of society: in modern societies where the state and the market provide many services, individuals tend to see their duties toward others as less important than they do in collectivist societies. This can be explained by the fact that in these modern societies, the state and the market provide alternative to mutual aid. In more traditional societies, people depend more mutual aid and therefore see themselves as having more more duties toward others (Baron & Miller, 2000; Fiske, 1992; Levine, Norenzayan, & Philbrick, 2001).

Mutual aid is no less constrained by fairness principles than strict reciprocity. If people want to treat others in a mutually advantageous way, they need to share the costs and benefits of mutual aid equitably. This means that the help given and the help received (when either situation arises) must be of comparable value. People who think that it is appropriate to hold the door for someone who is less than two meters away from it and who act accordingly have every reason to expect others to do the same—provided that they have equal chances to be the one holding the door (one cannot, for instance, ask people whose office is close to the door to always open the door for others). This means also that individuals can't ask others to do more than what they are ready to do themselves, and that they may depart from being seen as optimal partner if they provide help well beyond set mutual expectation, for instance by holding the door for someone who is 10 meters away, thereby sending the wrong signal that this is what they would expect when the roles are switched.

The amount of help we owe to each other depends also on the number of people involved in a particular situation. In a group of say ten friends, when one is in need, nine friends can help. When this is so, each must provide a 9th of the help needed. The smaller the group, the greater the duty of each towards a member in need, the larger the group, the lesser the duty. This group-size factor should play a role in explaining why people feel they have more duty toward their friends than toward their colleagues, toward their colleagues than toward their fellow citizens, and so on (Haidt & Baron, 1996).

2.3.3 Punishment: Proportionality between tort and compensation

The mutualistic account of punishment makes specific predictions. Indeed, to the extent that punishment is about restoring fairness, the more unfair the wrongdoing, the bigger the punishment should be. Anthropological observations have extensively shown that, in keeping with this prediction, the level of compensation in stateless societies is directly proportional to the harm done to the victim: for example, the wrongdoer owes more to the victim if he has killed a family member or eloped with a wife than if he has stolen animals or destroyed crops (Hoebel, 1954; Howell, 1954; Malinowski, 1926). Similarly, laboratory experiments have shown that in modern societies people have strong and consistent judgments that the wrongdoer should offer compensation equivalent to the harm inflicted to the victim or, if compensation is not possible, should incur a penalty proportionate to the harm done to the victim (Robinson & Kurzban, 2006).

On the other hand, taking the perspective of an altruistic model of morality, the function of punishment is to impose cooperation and deter people from cheating and causing harm. To that extent, punishment of a given type of crime should be calibrated so as to deter people from committing it. In many cases, altruistic deterrence and mutualistic retribution may favor similar punishments, making it impossible to directly identify the underlying moral intuition, let alone the evolved function. But in some cases, the two approaches result in different punishments. Consider

18

for instance two types of crime that cause the same harm to the victim and bring the same benefits to the culprit. From a mutualistic point of view, they should be equally punished. From an altruistic point of view, if one of the two types of otherwise equivalent crime is easier to commit, it calls for stronger deterrence and should be more heavily punished (Polinsky & Shavell, 2000; Posner, 1983). At present, we lack the large-scale cross-cultural experimental studies of people's intuitions on cases allowing clear comparisons to be able to ascertain the respective place of altruistic and mutualistic intuitions in matters of punishment (but see Baumard, 2011). What we are suggesting here is that, from an evolutionary point of view, not only should mutualistic intuitions regarding punishment be taken into consideration, but they may well play a central role.

2.4 Conclusion

The mutualistic approach not only provides a possible explanation of the evolution of morality, it also makes fine-grained predictions about the way individuals should tend to cooperate. It predicts a very specific pattern: Individuals should seek to make contributions and distributions in collective actions proportionate to each other; they should make their help proportionate to their capacity to effectively address needs; and they should make punishments proportionate to the corresponding crimes. These predictions match the particular pattern described by contractualist philosophers. Contractualist philosophers, however, faced a puzzle: They explained morality in terms of an implicit contract, but they could not account for its existence. A naturalist approach need not face the same problem. At the evolutionary level, the selective pressure exercised by the cooperation market has favored the evolution of a sense of fairness that motivates individuals to respect others' possessions, contributions and needs. At the psychological level, this sense of fairness leads humans to behave as if they were bound by a real contract.[10]

3. Explaining cooperative behavior in economic games

In recent years, economic games have become the main experimental tool to study cooperation. Hundreds of experiments with a variety of economic games all over the world have shown that, in industrialized as well as in small scale societies, participants' behavior is far from being purely selfish (Camerer, 2003; Henrich, J. et al., 2005), raising the question, If not selfish, then what? In this section, we investigate the extent to which the mutualistic approach to morality helps explain in a fine-grained manner this rich experimental evidence.

Here we consider only three games: the ultimatum game, the dictator game and the trust game. In the **ultimatum game**, two players are given the opportunity to share an endowment, say a sum of $10. One of the players, (the "proposer") is instructed to choose how much of this endowment to offer to the second player (the "responder"). The proposer can make only one offer that the responder can either accept or reject. If the responder accepts the offer, the money is shared accordingly. If the responder rejects the offer, neither player receives anything. The **dictator**

---

[10]'Some contractualist philosophers, such as David Gauthier (1986), explain the contractualist logic of moral decisions in terms of rational choice. Although this approach offers an ultimate explanation of our moral judgments, its proximal counterpart remains at odds with what we know about moral cognition: humans do not behave in a fair way because they have calculated that doing so is the most rational solution. In a way, however, the mutualistic theory can be seen as a translation of Gauthier's rationalistic theory into evolutionary and psychological terms.

game is a simplification of the ultimatum game. The first player (the "dictator") decides how much of the sum of money to keep. The second player (the "recipient"), whose role is entirely passive, receives the remainder of the sum. The **trust game** is an extension of the dictator game. The first player decide how much of the initial endowment to give to the second player with the added incentive that the amount she gives will be multiplied (typically doubled or trebled) by the experimenter, and that the second player, who is now in a position similar to that of the dictator in the dictator game, will have the possibility of giving back some of this money to the first player. These three games are typically played under conditions of strict anonymity (i.e., players don't know with whom they are paired and the experimenter does not know what individual players decided). Since the dictator game removes the strategic aspects found in the ultimatum game and in the trust game, it is often regarded as a better tool to study genuine cooperation and, for this reason, we will focus on it.

3.1 Participants' variable sense of entitlement

3.1.1 Cooperative games with a preliminary earning phase

In economic games, participants are given money but they may hold different views as the extent to which each has rights over this money. Do they, for instance, have equal rights, or does the player who proposes or decides how it should be share have greater rights? Rather than having to infer participants' sense of entitlement from their behavior, the games can be modified so as to give reasons to participants to see one of them as being more entitled to the money than the other. In some dictator games in particular, one of the participants – the dictator or the recipient – has the opportunity to earn the money that will be later allocated by the dictator. Results indicate that the participant who has earned the money is considered to have more rights over it.

In Cherry, Frykblom, & Shogren' (2002) study, half of the participants took a quiz and earned either $10 or $40, depending on how well they answered. In a second phase, these participants became dictators and were each told to divide the money they had earned between themselves and another participant who has not been given the opportunity to take the quiz. The baseline condition was an otherwise identical dictator game but without the earning phase. Dictators gave much less in the earning than in the baseline condition: 79% of the 10$ earners and 70% of the 40$ earners gave nothing at all, compared to 19% and 15% in the matching no-earning conditions. By simply manipulating the dictator's sense of entitlement, the transfer of resources is drastically reduced.

Cherry et al.'s study was symmetric to an earlier one of Ruffle's (1998). In Ruffle's study, it was the recipient who earned money by participating in a quiz contest and either winning the contest and earning $10 or losing and earning $4. That sum was then allocated by the dictator (who had not earned any money). In the baseline condition, the amount to be allocated, $10 or $4, was decided by the toss of a coin. Offers made to the winners of the contest were higher and offers made to the losers were lower than in the matching baseline conditions.

These two experiments suggest that participants attribute greater right to the player who has earned the money. When it is the dictator who has earned the money, she is less generous, and when it is the recipient who has earned the money, she is more generous than in the baseline condition. Having earned the money to be shared entitles the earner to a larger share, which is what a fairness account would predict.

A recent study of Oxoby and Spraggon (2008) provides a more detailed demonstration of the same kind of effects. In their study, individuals had the opportunity to earn money based on their performance in a 20 questions exam. Specifically, participants were given $10 (Canadian) by answering correctly between 0 and 8 questions; $20 by answering correctly between 9 and 14 questions; and $40 by answering correctly 15 or more questions. Three types of conditions were compared, conditions where the money to be allocated was earned by the dictators, conditions where it was earned by the recipients, and standard dictator game conditions where the amount of money was randomly assigned. In this last baseline conditions, on average, dictators allocated receivers 20 percent of the money, which is consistent with previous dictator game experiments. In conditions where the money was earned by the dictators themselves, they simply kept all of it (making, that is, the "zero offer" that rational choice theory predicts self-interested participants should make in all dictator games). In conditions where the money was earned by receivers, on the other hand, the dictators gave them on average more than 50 percent.

Oxoby and Spragon's study goes further in showing how the size of the recipients' earnings affects the way in which dictators allocate them. To recipients who had earned $40, no dictator made a zero offer (to be compared with 11% such offer in the corresponding baseline condition), and 63% of the dictators offered more than 50% of the money (to be compared with no such offer in the corresponding baseline condition). Offers made to recipients who had earned only the minimum of $10 were not statistically different from those made in the corresponding baseline condition. Offers made to recipients who had earned $20 were half-way between those made to the 40$ and the 10$ earners. Since $10 was guaranteed even to participants who failed answering any question in the quiz, participants could consider that true earnings corresponded to money generated over and above $10. As the authors note: "only when receivers earned CAN$ 20 or CAN$ 40 were dictators sure that receivers' property rights were not simply determined by the experimenter. These wealth levels provided dictators with evidence that these rights were legitimate in that the receiver had increased the wealth available for the dictator to allocate." The authors further note that "the modal offer is 50 percent for the CAN$ 20 wealth level and 75 percent for the CAN$ 40 wealth level, *exactly* the amount that the receiver earned over and above the CAN$ 10 allocated by the experimenter." In other words, the dictator gives the recipient full rights over the money clearly earned in the test. Overall, the authors conclude, such results are best explained in terms of fairness than in terms of welfare (e. g. "other-regarding preferences"). Dictators, it seems, give money to the recipients not in order to help them, but only because and to the extent that they think that they are entitled to it (see also Bardsley, 2008 for further experimental results).

3.1.2 The variability of rights explains the variability of distributions

The results of dictator game experiments with a first phase in which participants earn money suggest that dictators allocate money on the basis of considerations of rights. The dictator takes into account in a precise manner the rights both players may have over the money. In standard dictator games however, there is no single clear basis for attributing rights over the money to one or the other player and this may explain the variability of dictators decisions: Some consider they should give nothing, others consider they should give some money, and yet others consider they should split equally (Hagen & Hammerstein, 2006; for a similar point, see Heintz, 2005).

More specifically, there are three ways for participants to interpret standard cooperative games. First, some dictators may consider that, since the money has been provided by the experimenter without clear rationale or intent, both participants should have the same rights over it. Dictators thinking so would presumably split the money equally. Second, other dictators may consider that, since they have been given full control over the money, that they are fully entitled to keep it. After all, in everyday life, you are allowed to keep the money handed to you unless there are clear reasons why you are not. In the absence of evidence to the contrary, possession is commonly considered evidence of ownership. Dictators who keep all the money need not, therefore, be acting on purely selfish considerations. They may be considering what is fair and think that is fair for them to keep the money.[11] Third, dictators may consider that the recipient has some rights over the money—or else, why should have they been instructed to decide how much to give to the recipient?—, but that their different roles in the game is strong evidence that they have different entitlements. Dictators are in charge and hence can be seen as enjoying greater rights and as being fair in giving less than 50 percent to the recipient.

This interpretation of dictators' reasoning in standard versions of the game is confirmed by some of the first experiments on participants' sense of entitlement by Hoffman and Spitzer (1985) and Hoffman, McCabe, Shachat, & Smith (1996). They observe that when individuals must compete to earn the role of dictator, they give less to the recipient than they do in a control condition where they became dictator by, for instance, the flipping of a coin. In the same way, participants' behaviors vary whether the game is called *osotua* (a long-term relationship of mutual-help among the Maasai, Cronk, 2007) or *harambee* (a Kenyan tradition of community self-help events, Ensminger, 2004) or whether the game is framed as a community event or as an economic investment (Liberman, Samuels, & Ross, 2004; Pillutla & Chen, 1999). Participants use the name of the game to decide whether the money involved in the game belongs to them or is shared with the other participants.

There is an interesting asymmetry observed in games where participants' sense of entitlement is grounded on earnings or competition: dictators keep everything when they have earned the money, but do not give everything when it is the recipient who has earned the money. Why? Of course, it could be mere selfishness. More consistent with the detailed results of these experiments and their interpretation in terms of entitlement and fairness suggest, is the alternative hypothesis that dictators interpret their position as giving them more rights over the money than the recipient. Remember for instance that, in Oxoby and Spraggon's experiment, the modal offer is exactly the amount that the receiver earned over and above the $10 provided anyhow by the experimenter. In other words, dictators seem to consider both that they are entitled to keep the initial $10, and that the recipients are fully entitled to receive the money they earned over and above these $10.

The same approach can explain the variability of offers in ultimatum games. As Lesorogol writes:

"If player one perceives himself as having ownership rights over the stake, then (…) low offers would be acceptable to both giver and receiver. This would explain why many player

---

[11] Some participants may also think that there is no actual recipient. Therefore, it is not immoral to keep everything (see for instance Frohlich, Oppenheimer, & Kurki, 2004).

twos accepted low offers. On the other hand, if ownership is construed as joint, then (...) low offers would be more likely to be rejected as a violation of fairness norms, explaining why some players do reject offers up to fifty percent of the stake." (Lesorogol, forthcoming)

Explaining the variability of dictators' allocations in terms of the diverse manner in which they may understand their and the recipient's rights is directly relevant to explaining the variability of dictators' allocations observed in cross cultural studies. These behaviors correlate with local cooperative practices. In societies where much is held in common and sharing in a dominant form of economic interaction, participants behave as if they assumed that they have limited rights over the money they got from the experimenter. In societies where property rights are mostly individual and sharing is less common, dictators behave as if they assumed that the money is theirs. Take for instance the case of the Lamalera, one of the fifteen small scale society compared in Henrich, et al. (Henrich, J. et al., 2005):

> "Among the whale hunting peoples on the island of Lamalera (Indonesia), 63% of the proposers in the ultimatum game divided the pie equally, and most of those who did not, offered more than half (the mean offer was 58% of the pie). In real life, when a Lamalera whaling crew returns with a large catch, a designated person meticulously divides the prey into pre-designated parts allocated to the harpooner, crewmembers, and others participating in the hunt, as well as to the sailmaker, members of the hunters' corporate group, and other community members (who make no direct contribution to the hunt). Because the size of the pie in the Lamalera experiments was the equivalent of 10 days' wages, making an experimental offer in the UG [Ultimatum Game] may have seemed similar to dividing a whale" (p. 812)

Henrich et al. contrast the Lamalera to the Tsimane of Bolivia and the Machiguenga of Peru who "live in societies with little cooperation, sharing, or exchange beyond the family unit. (...) Consequently, it is not very surprising that in an anonymous interaction both groups made low UG offers." (p. 812) In accord with their cultural values and practices, Lamalera proposers in the ultimatum game think of the money as owned in common with the recipient, whereas Tsimane and Machigenga proposers see the money as their own and feel entitled to keep it.

To generalize, the inter-individual and cross-cultural variability observed in economic games may be precisely explained by assuming that participants aim at fair allocation and that what they judge fair varies with their understanding of the participants rights in the money to be allocated. The mutualistic hypothesis posits that humans are all equipped with the *same* sense of fairness but may distribute resources differently for at least two reasons:

1. They do not have the same beliefs about the situation. Remember for instance the differences between Europeans and Americans regarding the origin of poverty. Surveys indicate that Europeans generally think that the poor are exploited and trapped in poverty while Americans tend to believe that poor people are responsible for their situation and could get themselves out of poverty through effort (Note that both societies have approximately the same level of social mobility: Alesina & Glaeser, 2004).

2. They do not face the same situations (Baumard, Boyer, & Sperber, 2010). For instance, the very same good will be distributed differently if it has been produced individually or

collectively. In the first case, the producer will have a claim to a greater share of the good, while in the second the good will need to be shared amongst the various contributors. Whether the good is kept to one individual or shared between collaborators, the same sense of fairness will have been applied.

Such situational and informational variations may explain some cross-cultural differences in cooperative games. In the example above, Lamalera fishers give more than the Tsimane in the Ultimatum Game because they have more reason to believe that the money they have to distribute is a collective good. The Lamalera indeed produce most of their resources collectively, whereas the Tsimane produce their resources in individual gardens. Here, Lamalera and Tsimane do not differ in their preferences, and they all share the same sense of fairness, but because of differences in features of their everyday lives they do not frame the game in the same way.

Incidentally, children's beliefs may explain their behavior in economic games. Indeed, children younger than 7 seem to be shockingly ungenerous when playing these games (Bernhard, Fischbacher, & Fehr, 2006; Blake & Rand, 2010). Although these observations seem to suggest a late development of a sense of justice, it contrasts with other results in developmental psychology that demonstrate a very early emergence of a preference for helping rather than hindering behavior (Hamlin, Wynn, & Bloom, 2007), fairness-based behavior (Hamann, Warneken, Greenberg, & Tomasello, 2011; Warneken, Lohse, Melis, & Tomasello, 2011) and fairness-based judgments (Baumard, Chevallier, & Mascaro, submitted; Geraci & Surian, 2011; LoBue, Nishida, Chiong, DeLoache, & Haidt, 2011; McCrink, Bloom, & Santos, 2010; M. Schmidt & Sommerville, J., 2011). One way to reconcile these apparently contradictory findings starts from the observation that young children do not have the same experience or perspective as adults. While adults rarely if ever get money for free, receiving resources from other is actually the norm rather than the exception for children. Proposer might thus see themselves as fully entitled to the resource they get in the game, exactly as they are fully entitled to the candies or the toys given by their aunt or their older sibling. The apparent lack of generosity among children may have more to do with their understanding of the game than with a late development of their moral sense.

3.2. Exchanges

3.2.1 Proportionality between contributions and distributions

To the extent that the social selection approach is correct, considerations of fairness and impartiality should also explain the distribution of resources in cases where both participants have collaborated in producing them. As we have seen in section 2, the social selection approach predicts that the distribution should be proportionate to the contribution of each participant. This is of course not the only possible arrangement (Cappelen, Hole, Sorensen, & Tungodden, 2007). From a utilitarian point of view for instance, global welfare should be maximized; in the absence of relevant information, participants should assume that the rate of utility is the same for both of them; hence both participants should get the same share, whatever their individual contributions.

A number of experiments have studied the distribution of the money in situations of collaboration (Cappelen et al., 2007; Cappelen, Sorensen, & Tungodden, 2010; Frohlich, Oppenheimer, & Kurki, 2004; Jakiela, 2009; Jakiela & Berkeley, 2007; Konow, 2000). In Frohlich et al., (2004) for instance, the production phase involves both dictators and recipients proofreading a text

to correct spelling errors. One dollar of credit is allocated for each error properly corrected (and a dollar is removed for errors introduced). Dictators receive an envelope with dollars corresponding to the net errors corrected by the pair and a sheet indicating the proportion of errors corrected by the dictator and the recipient.

Frohlich et al. compare Fehr and Schmidt's (1999) influential "model of inequity aversion" with an expanded version of this model that takes into account "just desert". According to the original model, participants in economic games have two preferences, one for maximizing their own payoff, the other for minimizing unequal outcomes. It follows in particular from this model that proposers in the ultimatum game or dictators in the dictator game should never give more than half of the money, which would go both against their preference for maximizing money and their preference for minimizing equality. Frohlich et al. claim that people have also a preference for fair distributions based on each participant's contribution. This claim is confirmed by their results: the modal answer in their experiment is for participants to leave an amount of money exactly corresponding to the number of errors corrected by the recipient (30 subjects out of 73 are in this case). Contrary to the prediction following from Fehr and Schmidt's initial model, Frohlich et al. found that some of the dictators who had been less productive than their counterpart left more than 50% of the money jointly earned (8 dictators out of approximately 35 in this situation compared to none when the dictators had been more productive than their counterpart).

The pattern evidence in Frohlich et al. has also been found in experiments framed as transactions on the labour market. In Fehr, Gächter and Kirchsteiger (1997; see also Fehr, Kirchsteiger, & Riedl, 1993, 1998), a group of participants is divided into a small set of "employers" and a larger set of "employees." The rules of the game are as follows. The employer first offers a "contract" to employees specifying a wage and a desired amount of effort. The employee who agrees to these terms receives the wage and supplies an effort level, which need not equal the effort agreed in the contract. (Although subjects may play this game several times with different partners, each employer–employee interaction is an anonymous one-shot event.)

If employees are self-interested, they will choose to make no effort, no matter what wage is offered. Knowing this, employers will never pay more than the minimum necessary to get the employee to accept a contract. In fact, however, this self-interested outcome rarely occurs in the experiment and the more generous the employer's wage offer to the employee, the higher the effort provided. In effect, employers presumed the cooperative predispositions of the employees, making quite generous wage offers and receiving higher effort, as a means to increase both their own and the employee's payoff. More precisely, employees contributed in proportion to the wage proposed by their employer. Similar results have been observed in Fehr, Kirchsteiger, and Riedl (1993, 1998).

The trust game can also be used to study the effect of participants' contributions on the distribution of money. The money given by the first player to the second is usually multiplied by two or three. The total amount to be divided could therefore be seen as the product of a common effort of the two players, the first player being an investor who takes the risk of investing money, and the second player being a worker, who can both earn part of the money invested and return a benefit to the investor. Most experiments indeed report that Player 2 takes into account the amount sent by Player 1: The greater the investment, the greater the return (Camerer, 2003). Note moreover that

the more Player 1 invests, the bigger the risks she takes. Players Two aiming at a fair distribution should take this risk into account. This is exactly what Cronk et al. (2007; 2008) observed. In their experiments, the more Player 1 invests, the bigger not only the amount *but also the proportion* of the money she gets back (see also Willinger, Keser, Lohmann, & Usunier, 2003; and with a different result Berg, Dickhaut, & McCabe, 1995).

3.2.2 Talents and privileges

It is consistent with the mutualistic approach (according to which people behave *as if* had passed a contract) that, in a collective action, the benefits to which each participant is entitled should be a function of her contribution. How do people decide what count as contribution? This is not a simple matter. In political philosophy, for instance, the doctrine of choice egalitarianism defends the view that people should only be held responsible for their choices (Fleurbaey, 1998; Roemer, 1985). The allocation of benefits should not take into account talents and other assets that are beyond the scope of the agent's responsibility. In cooperative games, a reasonable interpretation of this fairness ideal would be to consider that a fair distribution is one that gives each person a share of the total income that equals her share of the total *effort* (rather than a share of the raw contribution). From the point of view of the social selection of partners, however, choice egalitarianism is not an optimal way to select partners: partners who contribute more be it thanks to greater efforts or to greater skills are more desirable as partners and hence their greater contribution should entitle them to greater benefits. Hence choice egalitarianism and partner-selection-based morality lead to subtly different predictions.

Cappelen et al. (2007) have tested these two types of prediction in a dictator game. In the production phase the players were randomly assigned one of two documents and asked to copy the text into a computer file. The value of their production depended on the price they were given for each correctly typed word (arbitrary rate of return), on the number of minutes they had decided to work to produce a correct document (effort), and on the number of correct words they were able to type per minute (talent). The question was: which factors would participants choose to reward? In line with choice egalitarianism and partner selection, almost 80% of the participants found it fair to reward people for their working time, that is, for choices that were fully within individual control (effort). Almost 80% of the participants found it unfair to reward people for features that were completely beyond their control (arbitrary rate of return). Finally, and more relevantly, almost 70% of the participants found it fair to reward productivity even if productivity may have been primarily outside individual control (talent). This confirms the predictions of partner selection.

The mutualistic approach thus predicts that people should be fully entitled to the product of their contribution. There are limits to this conclusion though: if what they bring to others has been stolen to someone, individual should not remunerate their contribution for it would mean being accomplice of the stealing. More generally, goods acquired in an unfair way are not gibe rights over the resources they help to produce. Cappelen et al. (2010) compared the allocation of money in economic games where the difference in input was either fair or unfair. At the beginning of the experiment, each participant was given 300 Norwegian Krone. In the production phase, participants were asked to decide how much of this money they wanted to invest, and were randomly assigned a low or a high rate of return. Participants with a low rate of return doubled their investment, while those with a high rate of return quadrupled their investment. In the distribution phase, two games

were played. Participants were paired with a player who had the same rate of return in one game and with a player who had a different rate of return in the other game. In each game, they were given information about the other participant's rate of return, investment level, and total contribution, and they were then asked to propose a distribution of the total income. The results show that the modal allocation decision is for participants to take into account the amount invested by each player but not the rate of return that differed in an unfair manner (43% of the participants were in line with this principle) (see also Burrows & Loomes, 1994 about effort and luck in a barganing game; Konow, 2000a for similar results with a benevolent third party).

This analysis might explain the developmental trends observed by Almås et al (2010). They observed a decline of egalitarian distribution during adolescence. This decline corresponds to an increasing awareness that some differences in productivity are due to pure luck and that other are under the control of individuals. At the beginning, children probably do not fully understand that some participants are more gifted than others and therefore they prefer an egalitarian distribution. As they come to understand that participants do not contribute equally to the common work, they realize that some participants deserve a bigger amount of money than other. The same moral logic may thus lead young children to be egalitarian and older children meritocratic.

3.3 Mutual aid

3.3.1 Rights and duties in mutual help

As we have seen in section 2, mutual aid works as a form of mutual insurance. Individuals offer their contribution (helping others) and get a benefit in exchange (being helped when they need it). A number of economic games have shown that indeed, people feel that they have the duty to help others in and of course, greater need calls for greater help need (Aguiar, Branas-Garza, & Miller, 2008; Branas-Garza, 2006; Eckel & Grossman, 1996).

When an economic game is understood in terms of mutual help this should alter participants decisions and expectations accordingly. Several cross-cultural experiments that frame economic games in locally relevant mutual help terms well illustrate this effect. Lesorogol (2007) for instance ran an experiment on gift giving among the Samburu of Kenya. She compared a standard dictator game with a condition where the players were asked to imagine that the money given to Player 1 represented a goat being slaughtered at home and that Player 2 arrived on the scene just when the meat was being divided. In the standard condition, the mean offer was 41,3% of the stake, (identical to a mean of 40% in a standard dictator game played in a different Samburu community; see Lesorogol, 2005). By contrast, the mean offer in the hospitality condition was 19.3%. Informal discussions and interviews in the weeks following the games revealed that in a number of real-world sharing contexts a share of 20% would be appropriate (Lesorogol, 2005). For instance, women often share sugar with friends and neighbours who request it. When asked how much sugar they would give to friends if they had a kilogram of sugar, most women responded that they would give a "glass" of sugar, about 200 grams.

Cronk (2007) compared, among the Maasai of Kenya, two versions of a modified trust games where both player were given an equal endowment (Barr, 2004). In one of the two versions, the game was introduced with the words "this is an *osotua* game" (as we mentioned above, in Maasai, an *osotua* relationship is a long-term relationship of mutual-help and gift-giving among two people).

Cronk observed that this *osotua* condition was associated with lower transfers by both players and with lower expected returns on the part of the first players. As Cronk explains, in an *osotua* relationship, the partners have a "mutual obligation to respond to one another's genuine needs, but only with what is genuinely needed." Since both players had received money, Player 2 was not in a situation of need and could not expect to be given much.

Understanding people's sense of rights and duties in mutualistic terms helps make sense of further aspects of Cronk's results. Compare a transfer of resources made in order to fulfill a duty to help the receiver with an equivalent transfer made in the absence of any such duty. This second situation is well illustrated by the case of an investor who lends money to a businessman. Since the businessman was not entitled to this money, he is indebted to the investor and will have to give her back a sum of money proportionate to her contribution to the joint venture. This corresponds to what we observe in the standard trust game. The more Player 1 invests, the more he gets back. By contrast, in a situation of mutual help, individuals do not have to give anything back in the short run (except, maybe to show their gratitude). What they provide in exchange for the help they enjoyed is an insurance of similar help should the occasion arise, a help the amount of which will be determined more by the needs of the person to be helped than by how much was received on a previous occasion.

Such an account of mutual help makes sense of Cronk's results. In his experiment, *osotua* framing was associated with a negative correlation between amounts given by the first player and amounts returned by the second. Player 2 returns less money to Player 1 in the context of mutual help than the context of investment. In the context of mutual help, Player 2 does not share the money according to each participant's contribution. She takes the money as a favour and gives only a small amount back as a token of gratitude. Participants reciprocate less in the mutual help than in the standard condition because they see themselves as entitled to the help they receive:

> "Although *osotua* involves a reciprocal obligation to help if asked to do so, actual *osotua* gifts are not necessarily reciprocal or even roughly equal over long periods of time. The flow of goods and services in a particular relationship might be mostly or entirely one way, if that is where the need is greatest. Not all gift giving involves or results in *osotua*. For example, some gift giving results instead in debt (*sile*). *Osotua* and debt are not at all the same. While [*osotua* partners] have an obligation to help each other in time of need, this is not at all the same as the debt one has when one has been lent something and must pay it back (see also Spencer,1965, p. 27)." (p. 353)

In this experiment, the standard trust game and the mutual help trust game exhibit two very different patterns. In the standard game, the more you give, the higher your rights on the money and the higher the amount of money you receive. In the mutual help game, the more you give to the other participant and the higher the amount of money she keeps. This contrast makes clear sense in a mutualistic morality of fairness and impartiality. Every gift creates an obligation. The character of the obligation however varies according to the kind of partnership involved. The resources you received may be interpreted as a contribution for a joint investment, and must be returned with a commensurate share of the benefits, or they may be interpreted as help received when you were entitled to it, and the duty of the to help when and the occasion arises and in a manner commensurate to the need of the person helped.

3.3.2 Refusals of high offers

A remarkable finding in cross-cultural research with the ultimatum game is that, in some societies, participants refuse very high offers (in contrast to the more common refusal of very low offers). Interpreting economic games in terms of a mutualistic morality suggests a way to explain such findings. Outside of mutual help, we claim, gifts received create a debt and a duty to reciprocate. Gifts, in other terms are not and are not seen as just altruistic. Of course, in an anonymous one-shot ultimatum game, reciprocation is not possible and there is no duty to do what cannot be done. But it is not that (or, arguably, not even possible) to shed off one's intuitive social and moral dispositions when participating in such a game, and to fully inhibit one's spontaneous attitudes to giving, helping, or receiving. Such inhibition should be even more difficult in a small traditional society where anonymous relationships are absent or very rare. Moreover, in some societies, the duty to reciprocate and the shame that may come with the failure to do are culturally highlighted. Gift giving and reciprocation are highly salient, often ritualized forms of interaction. From an anthropological point of view, it is not surprising therefore that the refusal of very high offers should have been particularly observed in small traditional New Guinean societies such as the Au and the Gnau where accepting a gift creates onerous debts and inferiority until the debt is repaid. In these societies, large gifts that it may be hard to reciprocate are often refused (Henrich, J. et al., 2005; Tracer, 2003). [12]

3.4 Punishment

3.4.1 Restoring fairness

Participants display a range of so called 'punishing' behaviors in economic games. Most of them, however, can be explain by direct self-interest. The ultimatum game, for instance, involves only two individuals. This is the kind of situation that triggers revenge behaviors because each partner has a direct interest in deterring cheating from the other (McCullough et al., 2010). For this reason, the game of choice for studying punishment has been the public good game (PGG). In a typical PGG, several players are given, say, 20 dollars each. The players may contribute part or all of their money to a common pool. The experimenter then triples the common pool and divides it equally among the players, irrespective of the amount of their individual contribution. A self-interested player should contribute nothing to the common pool while hoping to benefit from the contribution of others. Only a fraction of players however follow this selfish strategy. When the PGG is played for several rounds (the players being informed in advance of the number of rounds to be played), players typically begin by contributing on average about half of their endowment to the common pool. The level of contributions however decreases with each round, until, in the final rounds, most players are behaving in a self-interested manner (Ledyard, 1994). When the PGG is played repeatedly with the same partners, the level of contribution declines towards zero, with most players ending up refusing to contribute to the common pool (Andreoni, 1995; E. Fehr & Gächter, 2002). Further experiments have shown that, given the opportunity, participants are disposed to

---

[12] The fear of incurring a debt does not explain all refusals of very high offers. In other situations, the refusals seem to be motivated by the view that very high offers are unfair for the proposer (Bahry & Wilson, 2006; Hennig-Schmidt, Li, & Yang, 2006; Lesorogol, forthcoming).

punish others (i.e. to fine them) at a cost to themselves (Yamagishi, 1986). When, and only when such costly punishment is permitted, cooperation does not deteriorate.

Punishment is often seen as a fundamental way to sustain cooperation. In a mutualistic framework, however, the competition among partners for participation in cooperative ventures is supposed to be strong enough to select cooperative and indeed moral dispositions (Barclay, 2004, 2006; Barclay & Willer, 2007; Chiang, 2010; Coricelli et al., 2004; Ehrhart & Keser, 1999; Hardy & Van Vugt, 2006; Page et al., 2005; K. M. Sheldon et al., 2000; Sylwester & Roberts, 2010). Uncooperative individuals are not made to cooperate by being punished. Rather, they are excluded from cooperative ventures (an exclusion that is harming to them, and in that sense, can be seen a form of 'punishment' but that is not aimed at, and does not have the function of forcing them to cooperate).

Still, even in mutualistic interactions, punishment may be appropriate, but for other reasons.. First, although a public good game involves more than two individuals, the number of players is mall and each player may have an interest in incurring a cost to deter cheating. In average, revengeful individuals may end up being in more cooperative groups (McCullough et al., 2010). Second, as noted by Guala (2010), inflicting a cost is usually the only way for the participants to manifest their disappointment and it is clearly in their interest to warn their future partners that they are not going to accept further cheating. These self-serving motives can very well be combined with more moral motives. As we noted in section 2.2.3, it may indeed be morally required to help each others to fight against injustice (or, to put it differently, to refuse to be the accomplice of an immoral act). That's the reason why people feel compelled to support uprisings in dictatorships or to give money to human rights organization. Of course, this duty to punish is limited, exactly as is the duty to help others. Thus, most of us feel that they have a duty to send financial contribution to NGOs, but not to take up arms and risk their life to liberate a people. In economic games, the cost of punishing others is quite small (a couple of dollars) though, and punisher are usually involved in the interaction (they are thus not real third party and may have an interest in inflicting a cost to the cheater). Participants may thus feel that they have to spend their money to put an end to this unfair situation and to restore a fair balance among participants.

In such a perspective, punishment can be seen as a *negative distribution* aiming at correcting an earlier unfair positive distribution. If such is the goal of punishment, it should occur also in situations where there is no cooperation to sustain but where there has been an unfair distribution to redress.

Dawes et al. (2007) use a simple experimental design to examine whether individuals reduce or increase others' incomes when there is no cooperation to sustain. They call these behaviors 'taking' and 'giving' instead of 'punishment' and 'reward' to indicate that income alteration cannot change the behavior of the target and that none of the players did something worng. Participants are divided into groups of four anonymous members each. Each player receives a sum of money randomly generated by a computer; the distribution is thus arbitrary and to that extent unfair since lucky players do not deserve a larger amount of money than unlucky players. Players are shown the payoffs of other group members for that round and are then provided an opportunity to give 'negative' or 'positive' tokens to other players. Each negative token reduces the purchaser's payoff by one monetary unit (MU) and decreases the payoff of a targeted individual by three MUs; positive tokens decrease the purchaser's payoff by one MU and increase the targeted individual's payoff by

three MUs. Groups are randomized after each round to prevent reputation from influencing decisions and maintain strict anonymity.

The results show that players incurred costs in order to reduce or augment the income of other players even though this behavior plainly had no effect on what would happen in the subsequent rounds. Analyses show that participants were mainly motivated by considerations of fairness and impartiality, trying to achieve an equal division of wealth.[13] 68% of the players reduced another player's income at least once, 28% did so five times or more, and 6% did so ten times or more (out of fifteen possible times). Also, 74% of the players increased another player's income at least once, 33% did so five times or more, and 10% did so ten times or more.

Most (71%) negative tokens were given to above-average earners in each group, whereas most (62%) positive tokens were targeted at below-average earners in each group. Participants who earned ten MUs more than the group average received a mean of 8.9 negative tokens compared to 1.6 for those who earned at least ten MUs less than the group average. In contrast, participants who earned at least ten MUs less than the group average received a mean of 11.1 positive tokens (compared to 4 for those who earned ten MUs more than the group average). Overall, the distribution of punishment displays the logic of fairness: The more a participant received money, the more others would "tax" her. Conversely the less she received, the more she would get "compensated".

In an additional experiment, subjects were presented with hypothetical scenarios in which they encountered group members who obtained higher payoffs than they did. Subjects were asked to indicate on a seven-point scale whether they felt annoyed or angry (1, 'not at all'; 7, 'very') by the other individual. In the 'high inequality' scenario, subjects were told they encountered an individual whose payoff was considerably greater than their own. This scenario generated much annoyance: 75% of the subjects claimed to be at least somewhat annoyed, and 41% indicated to be angry. In the 'low-inequality' scenario, differences between subjects' incomes were smaller, and there was significantly less anger. Only 46% indicated they were annoyed and 27% indicated they were angry. Individuals apparently feel negative emotions towards high earners and the intensity of these emotions increases with income inequality. Moreover, these emotions seem to influence behaviour. Subjects who said they were at least somewhat annoyed or angry at the top earner in the high-inequality scenario spent 26% more to reduce above-average earners' incomes than subjects who said they were not annoyed or angry. These subjects also spent 70% more to increase below-average earners' incomes.

In another study, the same team examined the relation between the random inequality game and the PGG (Johnson, Dawes, Fowler, McElreath, & Smirnov, 2009). Participants played two games: A random income game measuring inequality aversion and a modified PGG with punishment. Their results suggest that those who exhibit stronger preferences for equality are more willing to punish free-riders in the PGG. The same subjects who assign negative tokens to high earners in the random income experiment also spend significantly more on punishment of low contributors in the

---

[13] To make sure that reciprocation was not a motivation, the authors conducted additional analyses. Results show that the number of negative tokens sent was not significantly affected by that negative tokens received in the previous round nor the number of positive tokens sent were not significantly affected by that of positive tokens received.

PGG[14] suggesting that even in this game punishment may well be not about only sustaining cooperation but about inequality.

In a replication (see supplementary material of Johnson et al. 2009), participants also had the opportunity to pay in order to help others and the results were nearly identical. Participants who, in the random income game reduced the income of high earners or increased that of low were more likely to punish low contributor in the PGG. These two studies are consistent with the fairness interpretation of punishment. At least some cases of punishment in PGGs are better explained as in terms of retribution than in terms of support to cooperation (see also Leibbrandt & López-Pérez, 2008 who show that third parties punish socially efficient but unfair allocations).

It could be granted that these results contribute to showing that equalitarianism is or can be a motivation in economic games, but leave open the question as to whether a preference for equality follows from a preference for fairness. After all, the notion of a fair distribution is open to a variety of interpretations. It might be argued that an unequal random distribution is not in itself unfair (since everybody's chances are the same), and therefore a preference for equality of resources may be seen as based on an equalitarian motivation more specific than, and independent from a general preference for fairness. If however humans' evolved sense of fairness is a proximal mechanism for social selection of desirable partners, then it can be given a more precise content that directly implies or at least favors equalitarianism in specific conditions. Given the choice to participate in a game with either an equal distribution of initial resources, or a random unequal distribution, most people being rationally risk-adverse, would, everything else being equal, choose the game with an equal distribution (except in special circumstances, for instance if this initial inequality provided a few of the partners with the means to invest important resources in a way that would end up being beneficial to all). Forced to play a game with an unequal and random allocation of initial resources but given the opportunity to choose their partners, most people would prefer partners whose behavior would diminish the inequality of the initial distribution. Being disposed to reduce inequality in such conditions is a desirable trait in cooperation partners. Hence fairness defined in terms of mutual advantage may or impartiality, in appropriate conditions, directly favor equalitarianism.

3.4.2 Explaining "anti-social" punishment

So-called anti-social punishment, that is, the punishment of people who particularly cooperative, has been observed in many studies and remains highly puzzling: Why do some participants punish those who give *more* than others to the common pool? In a recent study, Herrmann et al. (2008) ran a PGG with punishment in sixteen comparable participant pools around the world. They observed huge cross-societal variations. In some pools, participants punished high contributors as much as they punished low contributors, whereas in other pools, participants punished only low contributors. In some pools, antisocial punishment was strong enough to remove the cooperation-enhancing effect of punishment. Such behavior completely contradicts the view

---

[14] To be sure that envy was not a motivation, they compare the willingness to punish high earners in the random game when high earners are above the participants' income (envy) than above the group's average income (fairness). Analyses show that fairness does a much better job predicting punishment in the PGG. In particular, when the participant's own income is taken as a referent point, the relation between the willingness to punish high earners in the random game and the willingness to punish high earners in the PGGs ceases to be significant.

that the purpose of punishments is to sustain cooperation. Self-interested participants should neither contribute nor punish. Participants motivated to act so as to sustain cooperation should contribute and punish those who contribute less than average. By contrast, a mutualistic approach suggests a possible explanation for anti-social punishment.

Under what conditions might players consider that it is fair to punish high contributors? In PGG, participants have to decide the amount of money they want to give to the common pool. Let's assume that they want to contribute in a fair way. If so, by contributing to the common pool, they not only contribute to the common pool, they also indicate what they take to be a fair contribution. For the same reasons, they may view the contributions of others not just as money that will eventually be shared (and the more the better) but also as an indication of what others see as a fair contribution, and here they may disagree. When they find that a contribution lower than their own was unfairly low, they may blame the low contributor. Conversely, when they find that a contribution was unnecessarily high and much higher than their own, they may feel unfairly blamed, at least implicitly, by the high contributor. Moreover, if they are being punished by other players (and unless they are themselves high contributors), they have good reason to suspect that they are punished by people who contributed more than they did. If they feel that this punishment was unfair and deserves counter-punishment, then the obvious targets are the high contributors.

Herrmann et al.'s extensive study supports this interpretation. First, they observe, it is in groups where contributions are low that participants punish high contributors: The lower the mean contributions in a pool, the higher the level of antisocial punishment. Second, the participants who punish high contributors are those who gave small amounts in the first rounds, indicating thereby that they had low standards of cooperation from the start. Third, Herrmann et al. found that antisocial punishment increases as a function of the amount of punishment received, suggesting that, in such cases, it was indeed a reaction to what was felt to have been an unfair punishment for low but fair contribution. That they saw their low contribution as nevertheless fair and hence unfairly punished is evidenced by the fact that anti-social punishers did not increase their own level of contribution when they were punished for it. All these observations support an interpretation of antisocial punishment as guided by considerations of fairness (however misguided they may be).

Finally, Herrmann et al. found that norms of civic cooperation are negatively correlated with antisocial punishment. They constructed an index of civic cooperation from data taken from the World Values Survey and in particular from answers to questions on how justified people think tax evasion, benefit fraud, or dodging fares on public transport are. The more objectionable these behaviours are in the eyes of the average citizen, the higher the society's position in the index of civic cooperation. What they found is that antisocial punishment is harsher in societies with weak norms of civic cooperation. In these societies, people feel unfairly looked down upon by high contributors who expect too much from others. This observation fits nicely with qualitative research findings. For instance, in a recent article Gambetta and Origgi (2009) have described how Italian academics tacitly agree to deliver and receive low contributions in their collaborations and regard high contributors as cheaters who treat others unfairly by requiring too much of them.

Punishment, to conclude, may occur for a variety of reasons. Enforcement of cooperation is not the only possible reason, and need not be the main one. Even when the goal is to cause the other players to cooperate, this may be for selfish strategic reasons, thinking for instance that, in a

repeated PGG with only four participants, it is a good short-term investment to punish low cooperators and incite them to contribute to the common good (but see Falk, Fehr, & Fischbacher, 2005). There is evidence too that some participants punish both high and low contributors in order to increase their own relative pay-off, thus acting out of "spite" (Cinyabuguma, Page, & Putterman, 2004; Falk et al., 2005; Saijo & Nakamura, 1995). Still, what we hope to have shown is that, contrary to what is commonly supposed, a mutualistic approach can contribute to the interpretation of punishment and provide parsimonious fine-grained explanations of quite specific observations.

3.5 Conclusion

Experimental games are often seen as the hallmark of altruism. These games were originally invented by economists to debunk the assumption of selfish preferences in economic models. Since then, the debate has revolved around the opposition between cooperation and selfishness rather than on the logic of cooperation itself. Every game has been interpreted as evidence of cooperation or selfishness, and since altruism is the most obvious alternative to selfishness, cooperative games have been taken to favor altruistic theories (Gintis et al., 2003; Henrich, J. et al., 2005). In this article, we have proposed to explore another alternative to selfishness (mutualism) and to look more closely to the way participants depart from selfishness (through the moral parameters that impact on their decisions to transfer resources). Our hunch is thus that participants in economic games, despite their apparent altruism, are actually following a mutualistic strategy. When participants transfer resources, we argue, they do not *give* money (contrary to the appearances) they rather *refrain from stealing* money over which others have rights (which would amount of favoring one's side).

Because they were invented to study people's departure from selfishness rather than cooperation itself, classic experimental games however may not be the best tool to study the logic of human cooperation and test various evolutionary theories. Their very simple design, which was originally a virtue, turns out to be a problem (Guala & Mittone, 2010; Krupp et al., 2005; Kurzban, 2001; E. A. Smith, 2005; V. L. Smith, 2005; Sosis, 2005). Participants do not have enough information about the rights of each player over the money, they are blind to the rights, claims and entitlements that form the basis of cooperative decisions and need to fill the blank themselves, making the experiment very sensitive to all kinds of irrelevant cues and the results at odds with cooperative behaviors in real life (Chibnik, 2005; Gurven & Winking, 2008; Wiessner, 2009). These problems are not without solutions. As we have seen, the experimenter can fill the blanks (using a production phase or a real life story), making the interpretation of the game more straightforward, and allowing very precise hypotheses about contributions, property, gift, etc. to be tested. The future may lie in these more contextualized experiments which take into account that human cooperation don't just cooperate but cooperate in very quite specific ways.

4. Conclusion

The mutualistic theory of morality we propose is based on the idea that the evolution of human cooperation favoured, at the evolutionary level, mutually advantageous interactions that are sustained, at the psychological level, by a mutualistic morality. In this theory, we claim, the evolutionary mechanism (partner choice) leads precisely to the kind of behavior (fairness-based) that is observed in humans. This can be explained by the fact that the distribution of benefits in each interaction is constrained by the existence of outside opportunities, determined by the market of potential partners. In this market, individuals should never consent to enter into an interaction in

which the marginal benefit of their investment is lower than the average benefit they could receive elsewhere. If two individuals have the same average outside opportunities, they should both receive the same marginal benefit from each resource unit they invest in a joint cooperative venture. In the long run, we argued, such an evolutionary process should have led to the selection of a sense of fairness, a psychological device to treat other with impartiality.

Although individual selection is often thought to lead to a very narrow kind of morality, we suggested that partner selection can also lead to the emergence of a full-fledged moral sense that drives humans to be genuinely moral, to help each other and to demand the punishment of wrong-doers. This full-fledged moral sense may explain the kind of cooperative behavior observed in economic games such as the ultimatum game, the dictator game and the public good game. Indeed, in economic games, participants' behavior seem to aim at treating others in a fair way, distributing the benefit of cooperation according to individuals' contribution, taking others' claims to the resources into account, compensating them for previous misallocations, or sharing the costs of mutual help. In all these situations, participants act as if they had agreed on a contract or, as we claim, as if morality had evolved in a cooperative yet very competitive environment.

Of course, human cooperation is not exclusively guided by mutualistic norms. There are forms of cooperation where kin is favored over non-kin and in-group over out-group well beyond what considerations of fairness might sanction. As we pointed out, utilitarians favor acting for the greatest good of the greatest number even at the price of imposing unfair costs on specific individuals. While it is dubious that any human society has ever been governed by such utilitarian principles, individuals and groups have tried to live up to them. Various religious obligations that play an important role in human cooperation are not aimed at fairness and often conflict with it. There are legal norms that are often intended to be fair, but that, when there is a conflict between them and fairness are seen as taking precedence. This variety of norms, obligations, or preferences raises a terminological and two substantial issues.

The terminological issue has to do with the definition of morality. We have defined morality in terms of fairness (following a common tradition in ethics). It is possible of course to extend the notion of morality to a wider range of socially shared preferences that guide cooperation, but the price for this is giving up the intuition that an individual's moral norms should be consistent. More compelling and more substantial is the argument developed throughout this article that a specific and non-instrumental preference for fairness evolved as a distinct 'moral sense'. If you favor a more extensive definition of morality, call this a 'fairness sense' Even so, recognizing its very existence, however you call it, raises two substantial issues. First, how much human cooperative behavior is best explained in terms of this preference for fairness rather than in terms of other biologically or culturally evolved preferences? Regarding this first issue, we have made the case that considerations of fairness and impartiality provide uniquely fine-grained explanations of a great variety of experimental results and anthropological observations. In the future, experiments can and should be devised that test and possibly falsify predictions that are specific to the mutualistic approach, in particular when they differ from predictions entailed by other approaches. The second issue raised by the recognition of an evolved sense of fairness has to do with the way fairness norms and other norms of cooperation interact in biological and cultural evolution, in cognitive development and in behavior. Addressing this issue—which goes well beyond the scope of this article—cannot but be an

interdisciplinary effort recruiting evolutionary modeling, anthropological observations, and several branches of experimental psychology.

References

Adam, T. C. (2010). Competition encourages cooperation: client fish receive higher-quality service when cleaner fish compete. *Animal Behaviour*, *79*(6), 1183-1189.

Aguiar, F., Branas-Garza, P., & Miller, L. M. (2008). Moral distance in dictator games. *Judgment and Decision Making*, *3*(4), 344–354.

Aktipis, C. (2004). Know when to walk away: contingent movement and the evolution of cooperation. *Journal of Theoretical Biology*, *231*(2), 249-260.

Alesina, A., & Glaeser, E. (2004). *Fighting Poverty in the US and Europe: A World of Difference*. Oxford UK: Oxford University Press.

Alm\aas, I., Cappelen, A. W., SU00F8rensen, E. U., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176.

Alvard, M. (2004). Kinship, lineage identity, and an evolutionary perspective on the structure of cooperative big game hunting groups in Indonesia. *Human Nature*, *14*(2), 129-163.

Alvard, M., & Nolin, D. (2002). Rousseau's whale hunt? Coordination among big game hunters. *Current Anthropology*, *43*, 533.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256-274.

André, J. B., & Baumard, N. (2011). The evolution of fairness in a biological market. *Evolution*.

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *The American Economic Review*, 891-904.

Aspelin, P. (1979). Food distribution and social bonding among the Mamainde of Mato Grosso, Brazil. *Journal of Anthropological Research*, *J 35*, 309-27.

Aumann, R. J. (1981). Survey of repeated games. *Essays in game theory and mathematical economics in honor of Oskar Morgenstern*, *4*, 11–42.

Aumann, R. J., & Shapley, L. S. (1992). *Long-term competition: A game-theoretic analysis*. Dept. of Economics, University of California.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396.

Bahry, D., & Wilson, R. (2006). Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. *Journal of Economic Behavior and Organization*, *60*(1), 37-54.

Bailey, R. C. (1991). *The Behavioral Ecology of Efe Pygmy Men in the Ituri Forest, Zaire*. Ann Arbor: Museum of Anthropology, University of Michigan.

Balicki, A. (1970). The Netsilik Esquimo. Nueva York, Natural History Press.

Barclay, P. (2004). Trustworthiness and Competitive Altruism Can Also Solve the « Tragedy of the Commons ». *Evolution & Human Behavior*, *25*(4), 209-220.

Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, *27*, 325-344.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proc Biol Sci.*, *274*(1610), 749-53.

Barkow, J. (1992). Beneath New Culture Is Old Psychology: Gossip and Social Stratification. Dans J. Barkow, L. Cosmides, & J. Tooby (Éd.), *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press, USA.

Barnard, A., & Woodburn, J. (1988). Property, power, and ideology in hunter-gatherer societies: An introduction. *Hunters and Gatherers*, *2*, 4–31.

Baron, J., & Miller, J. (2000). Limiting the Scope of Moral Obligations to Help: A Cross-Cultural Investigation. *Journal of Cross-Cultural Psychology*, *31*(6), 703.

Barr, A. (2004). *10. Kinship, Familiarity, and Trust: An Experimental Investigation*. Foundations of Human Sociality. Oxford Scholarship Online Monographs.

Baumard, N. (2008). *Une théorie naturaliste et mutualiste de la morale*. Thèse de doctorat à l'Ecole des Hautes Etudes en Sciences Sociales, Philosophie et sciences sociales, Paris.

Baumard, N. (2010). Has punishment played a role in the evolution of cooperation? A critical review. *Mind and Society*.

Baumard, N. (2011). Punishment is a not group adaptation Humans punish to restore fairness rather than to help the group. *Mind and Society*, *10*(1).

Baumard, N., & Sperber, D. (submitted). Moral and reputation in an evolutionary perspective.

Baumard, N., Boyer, P., & Sperber, D. (2010). Evolution of Fairness: Cultural Variability. *Science*, *329*(5990), 388.

Baumard, N., Mascaro, O., & Chevallier, C. (in press).*Developmental Psychology*.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122-142.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912-915.

Bernstein, L. (1992). Opting out of the Legal System: Extralegal Contractual Relations in the Diamond Industry. *The Journal of Legal Studies*, *21*(1), 115-157.

Black, D. (2000). On the origin of morality. Dans L. Katz (Éd.), *Evolutionary origins of morality : cross-disciplinary perspectives* (p. xvi, 352 p.). Thorverton, UK ; Bowling Green, OH: Imprint Academic.

Blake, P. R., & Rand, D. G. (2010). Currency value moderates equity preference among young children. *Evolution and Human Behavior*, *31*(3), 210-218. doi:10.1016/j.evolhumbehav.2009.06.012

Boyd, R., & Richerson, P. (2005). Solving the Puzzle of Human Cooperation. Dans S. Levinson (Éd.), *Evolution and Culture* (p. 105-132). Cambridge MA: MIT Press.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. (2003). The evolution of altruistic punishment. *Proc Natl Acad Sci U S A.*, *100*(6), 3531-5.

Branas-Garza, P. (2006). Poverty in dictator games: Awakening solidarity. *Journal of Economic Behavior & Organization*, *60*(3), 306–320.

Brosig, J. (2002). Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior and Organization*, *47*(3), 275-290.

Broten, N. (2010). From Sickness to Death: The Financial Viability of the English Friendly Societies And Coming of the Old Age Pensions Act, 1875-1908.

Brown, W. M. (2003). Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm.

Bshary, R., & Grutter, A. (2005). Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Biology Letters*, *1*(4), 396.

Bshary, R., & Grutter, A. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, *441*(7096), 975-978.

Bshary, R., & Noë, R. (2003). The ubiquitous influence of partner choice on the dynamics of cleaner fish–client reef fish interactions. Dans P. Hammmerstein (Éd.), *Genetic and cultural evolution of cooperation* (p. 167–184). MIT Press.

Bull, J., & Rice, W. (1991). Distinguishing mechanisms for the evolution of co-operation. *Journal of Theoretical Biology*, *149*(1), 63.

Burrows, P., & Loomes, G. (1994). The impact of fairness on bargainin. *Empirical Economics*, *19*(2), 201-221.

Cadelina, R. V. (1982). Batak interhousehold food sharing: a systemic analysis of food management of marginal agriculturalists in the Philippines.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

Cappelen, A. W., Hole, A. D., Sorensen, E. O., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, *97*(3), 818-827.

Cappelen, A. W., Sorensen, E. O., & Tungodden, B. (2010). Responsibility for what? Fairness and individual responsibility. *European Economic Review*, *54*(3), 429–441.

Cashdan, E. (1980). Egalitarianism among Hunters and Gatherers. *American Anthropologist*, *82*(1), 116-120.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, *9*(2), 129–136.

Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review*, *92*(4), 1218-1221.

Chiang, Y. (2008). A Path Toward Fairness. *Rationality and Society*, *20*(2), 173.

Chiang, Y. (2010). Self-interested partner selection can lead to the emergence of fairness. *Evolution and Human Behavior*, *31*(4), 265-270.

Chibnik, M. (2005). Experimental economics in anthropology: A critical assessment. *American Ethnologist*, *32*(2), 198-209.

Cima, M., Tonnaer, F., & Hauser, M. (2010). Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience*.

Cinyabuguma, M., Page, T., & Putterman, L. (2004). On perverse and second-order punishment in public goods experiments with decentralized sanctioning. *Brown University, Department of Economics-Working Paper*.

Clark, M., & Jordan, S. (2002). Adherence to Communal Norms: What It Means, When It Occurs, and Some Thoughts on How It Develops. *New directions for child and adolescent development*, *95*, 3-25.

Clark, M., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, *37*(1), 12-24.

Clutton-Brock, T. (2002). Breeding Together: Kin Selection and Mutualism in Cooperative Vertebrates. *Science*, *296*(5565), 69-72.

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

Clutton-Brock, T., & Parker, G. (1995). Punishment in animal societies. *Nature*, *373*(6511), 209-216.

Coricelli, G., Fehr, D., & Fellner, G. (2004). Partner Selection in Public Goods Experiments. *Journal of Conflict Resolution*, *48*(3), 356-378.

Cronk, L. (2007). The influence of cultural framing on play in the trust game: A Maasai example. *Evolution and Human Behavior*, *28*(5), 352-358.

Cronk, L., & Wasielewski, H. (2008). An unfamiliar social norm rapidly produces framing effects in an economic game. *Journal of Evolutionary Psychology*, *6*(4), 283-308.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794-796.

DeScioli, P., & Kurzban, R. (2009). The Alliance Hypothesis for Human Friendship. (S. Pinker, Éd.)*PLoS ONE*, *4*(6), e5802. doi:10.1371/journal.pone.0005802

Dugatkin, L. (1995). Partner choice, game theory and social behavior. *Journal of Quantitative Anthropology*, *5*(1).

Dunbar, R. I. M. (1993). Co-evolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(4), 681-735.

Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, *16*, 181–191.

Ehrhart, K.-M., & Keser, C. (1999). Mobility and Cooperation: On the Run. *Série scientifique (CIRANO);99s-24*.

Elster, J. (2007). *Explaining social behavior: More nuts and bolts for the social sciences*. Cambridge ; New York: Cambridge University Press. Consulté de

Emery, G., & Emery, J. (1999). *A young man's benefit: The Independent Order of Odd Fellows and sickness insurance in the United States and Canada, 1860-1929*. McGill Queens Univ Pr.

Emlen, S. T. (1997). Predicting family dynamics in social vertebrates. *Behavioral ecology*, *4*, 228–253.

Ensminger, J. (2004). Market integration and fairness: evidence from ultimatum, dictator, and public goods experiments in East Africa. *Foundations of human sociality: economic experiments and ethnographic evidence from fifteen small-scale societies*, 356-381.

Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, 2017-2030.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137-140.

Fehr, E., & Henrich, J. (2003). Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. Dans P. Hammerstein (Éd.), *Genetic and cultural evolution of cooperation* (p. 55–82). Cambridge, Mass.: MIT Press.

Fehr, E., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica*, *65*(4), 833-860.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation*. *Quarterly journal of Economics*, *114*(3), 817-868.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics*, *108*(2), 437-459.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, *42*(1), 1-34.

Fiske, A. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *PSYCHOLOGICAL REVIEW-NEW YORK-*, *99*, 689-699.

Fleurbaey, M. (1998). Equality among responsible individuals. Dans J. F. Laslier (Éd.), *Freedom in Economics: New Perspectives in Normative Analysis* (p. 206-234). London: Routledge.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, *82*(2), 225-246.

Frank, R. (1988). *Passions within reason: The strategic role of the emotions* (Vol. 1). New York: Norton.

Frank, R., Gilovich, T., & Regan, D. (1993). The Evolution of One-Shot Cooperation: An Experiment. *ETHOLOGY AND SOCIOBIOLOGY*, *14*, 247-247.

Frohlich, N., Oppenheimer, J., & Kurki, A. (2004). Modeling other-regarding preferences and an experimental test. *Public Choice*, *119*(1), 91-117.

Fürer-Haimendorf, C. von. (1967). *Morals and merit: A study of values and social controls in South Asian societies*. London,: Weidenfeld & Nicolson.

Gambetta, D., & Origgi, G. (2009). L-worlds. The curious preference for low quality and its norms.

Gardner, A., & West, S. A. (2004). Cooperation and punishment, especially in humans. *The American Naturalist*, *164*(6), 753-764.

Gauthier, D. (1986). *Morals by agreement*. Oxford, New York: Clarendon Press ; Oxford University Press.

Geraci, A., & Surian, L. (s. d.). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153-172.

Gosden, P. (1961). *The friendly societies in England, 1815-1875*. Manchester University Press.

Grafen, A. (1990). Sexual selection unhandicapped by the Fisher process. *Journal of Theoretical Biology*, *144*(4), 473–516.

Greif, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review*, *83*(3), 525-548.

Guala, F. (2010). Reciprocity: Weak or Strong? What Punishment Experiments Do (and Do Not) Demonstrate. *Working paper*.

Guala, F., & Mittone, L. (2010). Paradigmatic Experiments: the Dictator Game. *Journal of Socio-Economics*, *39*(5), 578–584.

Gurven, M. (2004). To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, *27*.

Gurven, M., & Winking, J. (2008). Collective Action in Action: Prosocial Behavior in and out of the Laboratory. *American Anthropologist*, *110*(2), 179-190.

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theorical Popululation Biology*, *69*(3), 339-348.

Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, *316*(5827), 998.

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology*, *26*, 201-218.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331.

Hamilton, W. (1964). The genetical evolution of social behaviour I and II. *Journal of Theoretical Biology*, *7*, 1-16 and 17-52.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-9.

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, *32*(10), 1402.

Hare, R. (1993). *Without conscience: The disturbing world of the psychopaths among us*. New York: Pocket Books.

Heintz, C. (2005). The ecological rationality of strategic cognition. *Behavioral and Brain Science*, *28*(6), 825-6.

Hennig-Schmidt, H., Li, Z., & Yang, C. (2006). Why people reject advantageous offers—Non-monotonic strategies in ultimatum bargaining Evaluating a video experiment run in PR China. *Journal of Economic Behavior and Organization*.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., et al. (2005). « Economic man » in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Science*, *28*(6), 795-815; discussion 815-55.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., et al. (2006). Costly Punishment Across Human Societies. *Science*, *312*(5781), 1767-1770.

Henry, J. (1951). The Economics of Pilagá Food Distribution1. *American Anthropologist*, *53*(2), 187–219.

Herrmann, B., Gächter, S., & Thöni, C. (2008). Antisocial Punishment Across Societies. *Science*, *319*(1362).

Hobbes, T. (1651). *Leviathan, or, The matter, forme, & power of a common-wealth ecclesiasticall and civill*. London,: Printed for Andrew Ckooke [i.e. Crooke], at the Green Dragon in St. Pauls Church-yard.

Hoebel, E. A. (1954). *The law of primitive man; a study in comparative legal dynamics*. Cambridge: Harvard University Press.

Hoffman, E., & Spitzer, M. (1985). Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice. *Journal of Legal Studies*, *14*(2), 259-297.

Hoffman, E., McCabe, K., & Smith, V. (1996). Social Distance and Other-Regarding Behavior in Dictator Games. *The American Economic Review*, *86*(3), 653-660.

Howell, P. (1954). *A Manual of Nuer Law: Being an Account of Customary Law, Its Evolution and Development in the Courts Established by the Sudan Government*. Published for the International African Institute by the Oxford University Press.

Jakiela, P. (2009). *Equity vs. Efficiency vs. Self-Interest: On the Use of Dictator Games to Measure Distributional Preferences*. Working paper.

Jakiela, P., & Berkeley, U. C. (2007). How fair shares compare: Experimental evidence from two cultures. *Job Market Paper*.

Johnson, T., Dawes, C., Fowler, J., McElreath, R., & Smirnov, O. (2009). The role of egalitarian motives in altruistic punishment. *Economics Letters*, *102*(3), 192-194.

Kant, I. (1785). *Grounding for the metaphysics of morals; with, On a supposed right to lie because of philanthropic concerns*. Hackett Publishing Company.

Kaplan, H., & Gurven, M. (2001). The Natural History of Human Food Sharing and Cooperation: A Review and a New Multi-Individual Approach to the Negotiation of Norms. *Conference on the Structure and Evolution of Strong Reciprocity*. Santa Fe Institute.

Kaplan, H., & Hill, K. (1985). Hunting Ability and Reproductive Success Among Male Ache Foragers: Preliminary Results. *Current Anthropology*, *26*(1), 131-133.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 1072-1091.

Konow, J. (2001). Fair and square: the four sides of distributive justice. *Journal of Economic Behavior and Organization*, *46*(2), 137-164.

Konow, J. (2003). Which Is the Fairest One of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature*, *XLI*(December), 1188-1239.

Krebs, J. R., & Davies, N. B. (1993). *An introduction to behavioural ecology*. Wiley-Blackwell.

Krupp, D. B., Barclay, P., Daly, M., Kiyonari, T., Dingle, G., & Wilson, M. (2005). Let's add some psychology (and maybe even some evolution) to the mix. *Behavioral and Brain Sciences*, *28*(06), 828–829.

Kurzban, R. (2001). Are experimental economists behaviorists and is behaviorism for the birds? *Behavioral and Brain Sciences*, *24*(03), 420–421.

Kurzban, R., & DeScioli, P. (2008). Reciprocity in groups: Information-seeking in a public goods game. *European Journal of Social Psychology*, *38*(1), 139.

Landa, J. T. (1981). A Theory of the Ethnically Homogeneous Middleman Group: An Institutional Alternative to Contract Law. *The Journal of Legal Studies*, *10*(2), 349-362.

Ledyard, J. (1994). Public goods: A survey of experimental research. *Public Economics*.

Leibbrandt, A., & López-Pérez, R. (2008). *The envious punisher*. University of Zurich working paper.

Lesorogol, C. (forthcoming). Gifts or Entitlements: The Influence of Property Rights and Institutions for Third-party. Dans J. Henrich & J. Ensminger (Éd.), *Experimenting with Social Norms: Fairness and Punishment in Cross-Cultural Perspective*. New-York: Russell Sage.

Lesorogol, C. (2007). Bringing Norms In. *Current Anthropology*, *48*(6), 920-926.

Levine, R. V., Norenzayan, A., & Philbrick, K. (2001). Cross-cultural differences in helping strangers. *Journal of Cross-Cultural Psychology*, *32*(5), 543.

Liberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality and Social Psychology Bulletin*, *30*(9), 1175.

Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature*, *445*(7129), 727–731.

LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S., & Haidt, J. (2011). When Getting Something Good is Bad: Even Three-year-olds React to Inequality. *Social Development*.

Locke, J. (1689). *Two treatises of government*. London: Awnsham Churchill.

Malinowski, B. (1926). *Crime and custom in savage society*. New York: Harcourt, Brace & company.

Marlowe, F. (2009). Hadza Cooperation. *Human Nature*, *20*(4), 417-430.

Marshall, G., Swift, A., Routh, D., & Burgoyne, C. (1999). What Is and What Ought to Be: Popular Beliefs about Distributive Justice in Thirteen Countries. *European Sociological Review*, *15*(4), 349-367.

McAdams, R. (1997). The Origin, Development, and Regulation of Norms. *Michigan Law Review*, *96*(2), 338-433.

McCrink, K., Bloom, P., & Santos, L. R. (2010). Children's and adults' judgments of equitable resource distributions. *Developmental Science*, *13*(1), 37–45.

McCullough, M. E., Kurzban, R., Tabak, B. A., Shaver, I. P. R., & Mikulincer, M. (2010). Evolved mechanisms for revenge and forgiveness. *Understanding and reducing aggression, violence, and their consequences*, 221–239.

Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Science*, *18*(3).

Miller, W. (1990). *Bloodtaking and peacemaking : feud, law, and society in Saga Iceland*. Chicago: University of Chicago Press.

Nesse, R. (2007). Runaway social selection for displays of partner value and altruism. *Biological Theory*, *2*(2), 143-155.

Noë, R., van Schaik, C., & Van Hooff, J. (1991). The market effect: an explanation for pay-off asymmetries among collaborating animals. *Ethology*, *87*(1-2), 97-118.

Norton, M. I., & Ariely, D. (2011). Building a Better America—One Wealth Quintile at a Time. *Perspectives on Psychological Science*, *6*(1), 9.

Ohtsubo, Y., & Watanabe, E. (2008). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior*.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. The Political economy of institutions and decisions. Cambridge ; New York: Cambridge University Press.

Oxoby, R. J., & Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior and Organization*, *65*(3-4), 703-713.

Page, Talbot, Putterman, L., & Unel, B. (2005). Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency. *The Economic Journal*, *115*(506), 1032-1053.

Petersen, M. B., Sell, A., Toobyi, J., Cosmides, L., Petersen, M. B., Sell, A., Tooby, J., et al. (2010). and Criminal Justice: A Recalibrational Theory ofPunishment and Reconciliation.

Pillutla, M. M., & Chen, X. P. (1999). Social Norms and Cooperation in Social Dilemmas: The Effects of Context and Feedback. *Organizational Behavior and Human Decision Processes*, *78*(2), 81-103.

Polinsky, A. M., & Shavell, S. (2000). The Economic Theory of Public Enforcement of Law. *Journal of Economic Literature*, *38*(1), 45-76.

Posner, R. (1983). *The Economics of Justice*. Cambridge: Harvard University Press.

Pradel, J., Euler, H. A., & Fetchenhauer, D. (2008). Spotting altruistic dictator game players and mingling with them: the elective assortation of classmates. *Evolution and Human Behavior*.

Price, J. A. (1975). Sharing: The integration of intimate economies. *Anthropologica*, *17*, 3-27.

Rachlin, H., & Jones, B. A. (2008). Social discounting and delay discounting. *Journal of Behavioral Decision Making*, *21*(1), 29–43.

Raiffa, L., & Luce, R. D. (1957). *Games and decisions*. Wiley.

Ratnieks, F. L. W. (2006). The evolution of cooperation and altruism: the basic conditions are simple and well known. *Journal of Evolutionary Biology*, *19*(5), 1413–1414.

Rawls, J. (1971). *A theory of justice*. Cambridge, Mass.,: Belknap Press of Harvard University Press.

Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proc. R. Soc. Lond. B*, *265*, 427-431.

Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, *70*(4), 901–908.

Robinson, P., & Kurzban, R. (2006). Concordance and conflict in intuitions of justice. *Minn. L. Rev.*, *91*, 1829.

Rockenbach, B., & Milinski, M. (2011). To qualify as a social partner, humans hide severe punishment, although their observed cooperativeness is decisive. *Proceedings of the National Academy of Sciences*, *108*(45), 18307 -18312. doi:10.1073/pnas.1108996108

Roemer, J. (1985). Equality of talent. *Economics and Philosophy*, *1*(2), 151-181.

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, *23*(2), 247-265.

Sahlins, M. (1965). On the Sociology of Primitive Exchange in The Relevance of Models for Social Anthropology. Dans M. Banton (Éd.), *The Relevance of Models for Social Anthropology* (p. 139–236).

Saijo, T., & Nakamura, H. (1995). The« spite » dilemma in voluntary contribution mechanism experiments. *The Journal of Conflict Resolution*, *39*(3), 535-560.

Scanlon, T. (1998). *What we owe to each other*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge,: Harvard University Press.

Schmidt, M., & Sommerville, J. (2011). Fairness Expectations and Altruistic Sharing in 15-Month-Old Human Infants. *PLoS ONE*, *6*(10).

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*(35), 15073.

Sheldon, K. M., Sheldon, M. S., & Osbaldiston, R. (2000). Prosocial values and group assortation. *Human Nature*, *11*(4), 387-404.

Silberbauer, G. (1981). Hunter/gatherers of the Central Kalahari. *Omnivorous primates*, 455–498.

Smith, E. A. (2005). Making it real: interpreting economic experiments. *Behavioral and Brain Sciences*, *28*(06), 832–833.

Smith, V. L. (2005). Sociality and self interest. *Behavioral and Brain Sciences*, *28*(06), 833–834.

Sober, E., & Wilson, D. (1998). *Unto others : the evolution and psychology of unselfish behavior*. Cambridge, Mass.: Harvard University Press.

Sosis, R. (2005). Methods do matter: Variation in experimental methodologies and results. *Behavioral and Brain Sciences*, *28*(06), 834–835.

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*.

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853.

Tomasello, M., Melis, A., Tennie, C., Wyman, E., Herrmann, E., & Schneider, A. (in press). Two key steps in the evolution of human cooperation: The mutualism hypothesis.

Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. *Human morality and sociality: Evolutionary and comparative perspective*, 191–234.

Tooby, J., Cosmides, L., & Price, M. E. (2006). Cognitive adaptations for n-person exchange: The evolutionary roots of organizational behavior. *Managerial and Decision Economics*, *27*, 103-129.

Tracer, D. (2003). Selfishness and Fairness in Economic and Evolutionary Perspective: An Experimental Economic Study in Papua New Guinea. *Current Anthropology*, *44*(3), 432-438.

Trivers, R. (1971). Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, *46*, 35-57.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: the sequel. A kernel of truth in predictive cheating detection. *Evolution and Human Behavior*, *28*(4), 260-271.

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young Children Share the Spoils After Collaboration. *Psychological Science*, *22*(2), 267.

West-Eberhard, M. (1979). Sexual selection, social competition, and evolution. *Proceedings of the American Philosophical Society*, *123*(4), 222-234.

Wiessner, P. (1996). Leveling the hunter: constraints on the status quest in foraging societies. Dans P. Wiessner & W. Schiefenhövel (Éd.), *Food and the status quest* (p. 171-191). Oxford: Berghahn Books.

Wiessner, P. (2005). Norm Enforcement among the Ju/'hoansi Bushmen A Case of Strong Reciprocity? *Human Nature*, *16*(2), 115-145.

Wiessner, P. (2009). Experimental Games and Games of Life among the Ju/'hoan Bushmen= Les jeux expérimentaux et les jeux de la vie ches les Bochimans Ju/'hoan. *Current anthropology*, *50*(1), 133–138.

Willinger, M., Keser, C., Lohmann, C., & Usunier, J. (2003). A comparison of trust and reciprocity between France and Germany: experimental investigation based on the investment game. *Journal of Economic Psychology*, *24*(4), 447-466.

Woodburn, J. (1982). Egalitarian Societies. *Man*, *17*(3), 431-451.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*(1), 110-116.