

Ratio Working Paper No. 103

***The Puzzle of Altruism
Reconsidered: Biological
Theories of Altruism and One-
Shot Altruism***

Doron Shultziner
Arnon Dattner

The Puzzle of Altruistic Behavior Reconsidered: Biological Theories of Altruism and One-Shot Altruism

By

Doron Shultziner, Politics & IR Department, University of Oxford

Arnon Dattner, Department of Zoology, Tel Aviv University

Abstract

This paper critically examines the state of the literature in evolutionary biology regarding theories of altruistic behavior. The shared theoretical problems of Kin-selection and Group-selection are examined. Theoretical and severe methodological problems of Reciprocal Altruism theory are also discussed. We offer new conceptual clarifications of the Handicap Principle theory regarding costs and benefits to both the donor and the recipient of an altruistic act. We also summarize supportive empirical studies which demonstrate how Handicap Principle theory easily explains altruistic behavior on a different logic than the one employed by other theories of altruistic behavior. Finally, we discuss the phenomenon of one-shot altruism in order to evaluate, and distinguish between, the predictive and explanatory power of different theories of altruistic behavior.

• We thank Alan Grafen and Danny Priel for their help in preparing this paper. We are especially grateful to Avishag and Amotz Zahavi for their valuable comments. This paper was presented at the “Trust, Reciprocity and Social Capital: The 2006 Ratio Colloquium for Young Social Scientists” Stockholm, 24th-26th of August 2006. We thank Niclas Berggren and Andreas Bergh of the Ratio Institute for organizing and sponsoring the colloquium which gave us the incentive to write this paper.

The puzzle of altruistic behavior among human beings and other species has puzzled biologists ever since Darwin [1859]. The puzzle is about explaining how an apparent altruistic behavior may emerge in an evolutionary process which is understood as a continual struggle for existence and the survival of the fittest. Altruistic behavior is an act that advances the fitness of the recipient (or recipients) of the altruistic act and at the same time incurs fitness costs on the donor of the altruistic act. Biologists assume that for such a behavior to evolve it must confer some important direct or indirect benefits on the donor, otherwise this behavioral tendency would not endure the process of natural selection.

Several biological theories have been proposed to resolve this puzzle of altruism. The first three theories in chronological order are ‘group selection’ (GS); ‘kin selection’ (KS); and ‘reciprocal altruism’ (RA). Despite these theoretical efforts, there is great confusion and disagreements in the biological literature about and between these alternative explanations.

The contribution of this paper to the study of altruism and cooperation is threefold. First we critically examine the state of the literature on altruistic behavior and point to some fatal flaws of theoretical inference that have not received proper attention in the literature so far. Second, we present Zahavi’s (1975, 1977a) ‘handicap principle’ (HP) as an alternative theory to altruistic behavior. We make important conceptual clarifications of HP theory that we hope will contribute to its understanding in general and to the understanding of altruistic behavior in particular. Third, we discuss the interesting phenomenon of one-shot altruism as a key to evaluate the explanatory and predictive power of the different theories.

The paper proceeds as follows. In section 1, we offer a critical assessment of two theories of altruistic behavior: GS and KS. Group selection and kin selection share similar theoretical flaws which we discuss. Some supporters of GS have been arguing that due to the similarities between KS and GS, GS models and arguments are validated and therefore become legitimate explanations of altruism. We argue to the contrary, namely, that due to their similar theoretical flaws both GS and KS should be questioned and modified.

In section 2, we present the Handicap Principle theory and advance some conceptual clarifications that have not yet been made explicit in the literature. First, we discuss the costs and benefits to the altruist’s fitness by distinguishing between

costs to survival (or viability) and benefits to reproduction. This simple explicit distinction may clarify many previous misunderstandings of this theory. Secondly, and with relation to the previous distinction, we emphasize the underlying competitive nature of altruistic behavior and its potential fitness costs to the recipient.

In section 3, we present and critically examine reciprocal altruism theory. We discuss some major theoretical problems in the ways this theory has been stretched, modified and tested. We review empirical works that have been widely cited as supporting RA and we contend that these empirical studies are highly problematic and do not directly support RA theory. We also explain how some anomalies in RA interpretations can be explained according to HP.

In section 4, we discuss altruism among humans and the phenomenon of one-shot altruism. This type of behavior poses interesting explanatory challenges to biological theories of altruism towards non-kin and to RA in particular. We compare the possible interpretations of RA and HP to one-shot altruism and we propose why this special case of altruistic behavior can further illuminate the altruism puzzle.

1. Group Selection and Kin Selection

This section explores the logic of GS and KS, the specific problems of each theory and their shared theoretical flaws.

1.1 Group Selection and Altruism

Group selection theory, in its initial and essential form, proposes that individual traits, such as altruistic behavior, have prospered *because* they confer group benefits. According to GS, altruistic behavior does not benefit the altruistic individual directly. Groups with altruistic individuals have better survival chances in the competition against other less altruistic groups, and hence altruistic alleles can prosper in the population. This type of basic logic is shared by all group selection models, including later models with more sophisticated mathematical formulations (compare Wynne-Edwards 1962; Wilson 1977; Boyd and Richerson 1990, 1992; Sober and Wilson 2000; Gintis 2000; Gintis et. al. 2001, 2003).

GS models employ a reverse reasoning in explaining altruistic behavior, namely, group benefits dictate and explain individual behavior, not vice versa. Individual benefits are an indirect outcome of group benefits, whereas altruistic behavior is costly to one's fitness in a direct sense. GS models have been severely criticized for this type of logic (Williams 1966; Maynard-Smith 1964, 1976) and consequently fell to disrepute. The main and persistent problem with GS is that it

cannot explain why social parasites will not exploit the group and its altruists. Social parasites will be better-off defecting and not contributing to the group and they will prosper at the expense of altruistic individuals in the group. Hence, the system is bound to collapse and to lose its altruistic alleles.

Despite these great theoretical difficulties, group selection arguments are still prevalent in the biological literature (e.g. Boyd and Richerson 1990, 1992; Sober and Wilson 2000; Gintis 2000; Gintis et. al. 2001, 2003). One line of GS reasoning is that punishment provides a solution to social parasites. GS advocates contend that if punishment is allowed into the equation then the problem of free-riding is solved and cooperation can be sustained at high levels (Boyd and Richerdson 1992; Gintis 2000). However, as Fehr and Gächter (2002: 137) argue, “this ‘solution’ begs the question of who will bear the cost of punishing the free riders. Every individual in the group will be better off if free riding is deterred, but nobody has an incentive to punish the free riders.” Fehr and Gächter frame this as a ‘second-order public good’ problem and go on to examine the amount of punishers necessary to maintain cooperation. Interesting as this question may be, it nevertheless leaves the puzzle of altruism unsolved because punishment is an altruistic act in itself and explaining one altruistic act with another altruistic act is a circular argument. That is to say, if punishment is required to make altruism work then this poses a further restriction on the evolution of altruism due to the multiple costs entailed both in altruistically cooperating and in altruistically punishing.

1.2. Kin Selection and altruism

Kin selection theory proposes that altruistic behavior evolved because of its benefits to the inclusive fitness of the carrier of the altruistic alleles (Hamilton 1964). Altruistic behavior can evolve and become widespread if the genetic relatedness of the recipient (R) multiplied by the benefit received from the altruistic act (B) is bigger than the costs to the donor of the altruistic act (C): $R \times B > C$. Namely, an altruistic individual can increase his or her inclusive fitness by helping kin. By behaving altruistically towards siblings (who share 50 percent of their genetic material), an individual can be seen as spreading his or her own genes in the population, including the alleles that encode for the altruistic behavior. Altruistic behavior towards kin, thus, is actually an indirect investment in, or defense of, one’s own genes. Accordingly, fitness is measured not only by one’s own survival and reproduction but also by the survival and reproductive success of one’s kin. This logic predicts that altruism would be strongest between parents and their offspring,

between siblings, and decreasingly to more distant kin until the relatedness is negligible.

Compelling as this theory may be in first appearance, it has some obvious shortcomings in explaining altruistic behavior. To begin with, it is evident that there are various manifestations of altruistic behavior between non-kin and other unrelated individuals, which violate Hamilton's rule of $R \times B > C$. Examples are food sharing among chimpanzees, mobbing predators, allofeeding and adoption among social birds, as well as incubating on other non-related bird's eggs, assisting in raising other's offspring among social birds and primates, blood-sharing in vampires, and many other examples. The theory of kin-selection is unable to explain such altruistic behavior due to the crucial factor and prerequisite of sufficient genetic relatedness between the donor and the receiver of an altruistic act.

More seriously, though, is a theoretical shortcoming that has been, paradoxically, pointed out by GS theorists in order to support GS reasoning. Sober and Wilson (2000: 191) make the observation that "kin selection is a kind of group selection, in which the groups are composed of relatives." Similarly, Zahavi and Zahavi (1997) claim that KS is GS among kin. Altruism among siblings and more distant related kin is inconsistent with individual-selection reasoning because "When an altruistic individual helps a related individual who is selfish, the donor still has a lower fitness than the recipient. The fact that they are related does not cancel this fundamental fact" (Sober and Wilson 2000: 191-192). That is, individuals that free-ride on their siblings' altruism can flourish at the latter's expense and hence the system becomes unstable, similar to problems of altruisms and cooperation in groups of not closely related individuals (see also Zahavi 1995, 2003).

Interestingly, GS theorists make the opposite claim that the similar logic of KS and GS strengthens their argument about the validity of GS models. Like Zahavi (1995, 1997, 2003), however, we claim to the contrary, namely, that GS and KS are equally problematic in explaining the evolution and logic of altruistic behavior. And in fact, if the postulated explanatory factor of KS is prone to the same theoretical flaws as in GS models, then "It is certainly not reasonable to reject GS and accept KS" (Zahavi 1995: 2). Hence, we argue that KS should be modified or abandoned altogether as a theory of altruism.

Whereas it is clear that evolution works at the genetic level and that selection ultimately occurs between different alleles, it nevertheless seems theoretically unsound to assume that genetic relatedness is also the level of selection

other than individual selection, assuming that the two can be differentiated. Evolution does ultimately affect the genetic level, but the basic unit of selection for the explanation of how this genetic material is spread or lost and of how altruistic alleles evolve is best understood and explained on an individual-selection level and not on its genetic expression and outcomes.

2. The Handicap Principle

Chronologically we should have now discussed the theory of reciprocal-altruism. However, in order to assess this influential theory we first need to present 'handicap principle' (HP) theory.

HP was first presented by Amotz Zahavi in 1975 in relation to mate-selection (or sexual selection) and was applied by Zahavi to the question of altruism already in 1977. The theory has suffered from disrepute and misunderstandings until the late 1980's. The situation changed, however, with Alan Grafen's (1990a, 1990b) two influential papers that gave a mathematical expression to the handicap principle.¹ Ever since the publication of Grafen's papers, Zahavi's handicap principle has gained increasing recognition and application in different fields of scholarship. Nevertheless, peculiarly, many biologists accept HP as an explanation to costly signaling phenomena but fail to apply the principle explicitly to altruism, although Zahavi (1977a) did make that connection explicit already three decades ago (see also Roberts 1998). Let us, then, recap the logic of HP and reemphasize why it applies equally to the explanation of altruism. By way of reemphasizing why and how HP equally applies to the phenomenon of altruism we also offer some conceptual clarifications of the theory regarding costs to fitness. We hope these clarifications will make HP theory more accessible and understandable to skeptics (Wright 1999).²

HP is an individual-selection theory and as such its basic logic follows the assumption that altruistic behavior must be beneficial primarily to the *individual's* overall fitness, not necessarily to his group or kin. Benefits to kin and the group are possible and quite common *outcomes* of altruistic behavior, but not the prime causes of the individual's behavior. HP has three intertwined assumptions that constitute the theoretical apparatus.

¹ In the first paper Grafen concludes that "the main biological conclusions of this paper are the same as those of Zahavi's original papers on the handicap principle" (1990a: 487) and in the second paper Grafen adds that "These verbal arguments are really just as convincing as all the mathematics, and their language makes clear the strong connection with Zahavi's arguments. This shows that the models given in this paper are really models of Zahavi's handicap principle" (Grafen 1990b: 541). Compare also to Zahavi's (2003) reflections on these issues.

² Still, these conceptual clarifications are not modifications of the theory or supplements to it but rather restatements of the same logic in more suitable and accurate concepts.

The first initial assumption of HP pertains to *phenotypic differences among individuals*. Phenotypic differences are – but not always – a result of genetic variability.³

The second assumption of HP is that altruistic behavior constitutes *a costly act of communication*. Individuals may look morphologically similar but will nevertheless differ in important respects regarding their qualities. The information about an individual's qualities is important to potential mates, partners and rivals. Females, for example, will prefer high-quality males because this choice would increase their own fitness. Males, on the other hand, benefit from advertising their qualities to attract females. Individuals that compete over rank and status also benefit from acquiring knowledge about their rival's qualities because this type of information saves time, energy and may reduce the risk of injury involved in violent ways of pursuing this essential information.

In order for this system to be functional, however, communication must be reliable. Communication becomes meaningless if it does not reliably correspond to the qualities being advertised, e.g. by cheating. In addition, if signals become equally costly to (or as easily produced by) individuals of different quality then the signals lose their meaning and may be selected-out (Zahavi and Zahavi 1997: chapter 4). Therefore, HP proposes that reliable communication requires handicaps or costs: “a signal is reliable when the difficulty of its performance is related to its meaning in quantity and quality” (Zahavi 1977a: 254). Communication which lacks costly signaling is not honest and therefore becomes prone to cheaters.

Furthermore, HP is the only theory that proposes a logical connection between the costly signal and its meaning (Zahavi 1977, 1997). Namely, handicaps are not random; they convey a message particular to the trait or traits being advertised. For example, individuals who handicap themselves by donating food are signaling that they are good foragers and individuals who engage in costly physical displays signal their strength, etc.

The third assumption pertains to *differential costs* involved in costly signaling. Given the first assumption of phenotypic variability, and supplemented with the second assumption that honest communication requires costs, a third assumption logically follows from the previous two assumptions that the costs of signaling differ

³ Due to genetic variability (caused by meiosis and mutation) there are almost no two individuals with a similar genotype (with the exception of identical twins). Different genotypes give rise to different phenotypes and this effect can become more pronounced given environmental factors (e.g. differential nutrition or parental care, random causes like injuries, etc.).

between individuals. Reliable communication is enabled and ensured *because* the costs of honest signaling differ between individuals in the following two main ways.

First, the higher the quality of the individual, the less costly the altruistic behavior is. It is easier to an individual of higher quality to bear the *same* type or degree of a costly behavior relative to an individual of lower quality. The lower quality individual may be able to carry the costs associated with this behavior only with graver costs. For instance, an individual that is not a good hunter or forager will find it harder to bear the costs of altruistic acts such as feeding other individuals, serving as a sentinel, mobbing a predator, etc.

Secondly, the higher the quality of the individual, the more costs the individual can endure. That is, not only do similar costs differ with regard to their impact on an individual's survival, but also the degree or amount of costly behavior that individuals can endure varies. Higher quality individuals are capable of performing types or degree of costly behavior that lower quality individuals are incapable of. The ability to endure certain types or degrees of handicaps signals-out those individuals that are more fit than those who are less fit.

These three assumptions are the pillars on which HP stands. HP theory, however, suffered from disrepute and many biologists were not persuaded by it. Once a mathematical expression and proof⁴ to HP was offered by Grafen (1990a, 1990b), HP began to gain more attention and recognition.⁵ The persuasiveness of mathematics, however, was not the only reason that kept HP from being applied in biology and related fields. HP appeared in the eyes of many scholars as counterintuitive or even absurd (compare Grafen 1990b: 527). It was hard to convince scholars that altruistic behavior, for instance, will evolve not in spite, but *because*, it is costly. It seems counter intuitive to think that a handicapped individual will do better than a non-handicapped individual in a process of competitive natural selection.

⁴ Mathematical proofs are more highly regarded among biological theoreticians than are verbal proof and argumentation. Mathematical proof, however, does not have a higher epistemological status than logical argumentation. In fact, every theoretical argumentation can be represented in mathematical or other logical formulations. False theories can nevertheless be represented in mathematical models which do not explain the phenomenon they seek to explain. Hence, several theoretical models, backed up by different mathematical models, can explain the same phenomenon. A theory, therefore, should be evaluated more by its logic and explanatory power rather than by the sophistication of its mathematical or game theoretic models. Zahavi reflects upon these issues in his later papers (Zahavi 1999, 2003).

⁵ For example, John Maynard Smith (1985) who denied the theory up into the mid-1980's wrote his own mathematical formula to HP (Maynard Smith 1991) after Grafen's papers were published and resonated in the field of theoretical biology.

Nevertheless, this point is not paradoxical at all when it is presented in suitable concepts. The best way to resolve this paradox is to consider and present more accurately the costs and benefits to an altruist's fitness by the two main components of fitness: survival (or viability) and reproduction. Of the two, reproduction is the more important because, at the end of the day, an individual's fitness is commonly measured by the number of live offspring that the individual has. It is not uncommon in nature that individuals employ a strategy of enduring extreme costs to their survival in exchange for reproductive success. In fact, most or all parental-care cases are examples of survival costs exchanged with propagation benefits. With regard to the phenomenon of altruistic behavior, the costs to the donor are primarily in terms of survival (investing energy and time in others, food sharing, risk taking) and the benefits are in terms of higher reproductive chances.

Altruistic behavior is a form of investment (costly communication) in advertising the individual's qualities, thereby increasing or sustaining the altruist's social prestige, rank and status (Zahavi 1995: 2). Social prestige and status benefit the altruist by deterring rivals and improving the altruist's social standing in the group.⁶ It is well known that 'social prestige' and high rank and status are key factors in reproduction. Individuals of higher rank tend to reproduce more and their offspring tend to have higher survivorship rates (Barkow 1989; de Waal 1982; Grammer 1996; Hold-Cavell 1996; Pusey, Williams and Goodall 1997; Turke and Betzig 1985). In this perspective, costs to survival can also be seen as an investment, an advertising effort on the part of donors that serves to increase their own fitness (see also Zahavi 2003).

A final clarification pertains to the effects of altruistic behavior on the recipient's fitness, especially in cases of intra-sexual competition. Since altruism is a competitive strategy, it follows that there may also be some costs involved in being the recipient of an altruistic act (see also Roberts 1998). A recipient may benefit in terms of food or energy, but the recipient's chances of reproduction may be offset because an act of receiving emphasizes the recipient's weakness and inferiority in comparison to the donor. For this reason, individuals often refuse to be recipients of altruistic acts and may even aggressively react to donors (see below).

⁶ Shifts in power relations within the group are not always reflected in a change in status or rank but can manifest themselves also in the degree that a dominant individual may need to compromise with subordinates, see (Kaloshov, Zahavi and Zahavi 2005; Zahavi 1995, 1997).

2.1 Supporting Empirical Work of the Handicap Principle and Altruistic Behavior

Zahavi's work with Arabian Babblers (*Turdoides squamiceps*), a group-breeding songbird, demonstrates several key issues in the nature of altruism. Amotz Zahavi with his wife Avishag Zahavi, and several other students who have worked with them over more than three decades at the Hatzeva nature reserve in southern Israel, have made close observations of the babblers while taking into account the intricacies of group and inter-group composition. Their studies have shown that babblers actually compete over performing altruistic acts in several respects.

First, the Zahavis showed that higher-ranking males tend to perform substantially more altruistic acts than lower-ranking individuals do. On the other hand, high-ranking individuals receive the least donations, if at all. For instance, the alpha male serves as a sentinel significantly more than other high ranking individuals in the group and mobs predators closer than other members do (Zahavi and Zahavi 1997). Generally, the higher the rank is, the more such altruistic acts that the babbler performs.

Second, the Zahavis showed that altruistic behavior is almost exclusively unidirectional from a dominant individual to a lower ranking individual. Moreover, altruistic behavior is not randomly performed towards any lower ranking individual, nor is it performed towards individuals who may benefit the most from the altruistic act. Rather, altruistic acts are directed towards individuals that are closest to the donor in rank. This pattern of a unidirectional and intentional altruistic behavior towards particular individuals is also exemplified by the fact that dominants often ignore a subordinate individual that begs for food while the dominants are on their way to feed a closer ranked subordinate (Kalishov, Zahavi and Zahavi 2005).

Third, it was also shown that individuals often refuse to be recipients of altruistic acts, such as being replaced as a sentinel, being helped at the nest and being fed (Carlisle and Zahavi 1986; Kalishov, Zahavi and Zahavi 2005; Zahavi and Zahavi 1997). Refusals to being the recipient of altruistic acts are common among high-ranking as well as low-ranking individuals. The alpha male is particularly resistant to be a recipient of altruistic acts.⁷

⁷ The rare cases in which subordinate babblers did feed dominants were in small groups surrounded by strong groups. In these situations the dominant male is much more dependent upon the help and cooperation of the subordinate male and thus makes concessions to the subordinate both in being willing to be a recipient of food and in conceding reproduction opportunities to the subordinate (see also Kalishov, Zahavi and Zahavi 2005).

Forth, it was shown that babblers tend to be aggressive and violent in response to altruistic acts (Carlisle and Zahavi 1986; Kalishov, Zahavi and Zahavi 2005). High-ranking babblers tend to be violent toward individuals of lower rank that attempt to feed them. In several cases, a dominant individual has aggressively responded to a lower ranking individual that tried to feed him or her, snatched the food from the mouth of the subordinate, and coerced the subordinate into becoming the recipient by forcing the food into the subordinate's throat. Alpha males are especially sensitive to such attempts. In fact, in the rare occasions observed in which a subordinate individual has attempted to feed an alpha male, this behavior has triggered an outbreak of violence between the two, leading to the killing or escape of one of the individuals from the group. Moreover, dominant babblers tend to punish lower ranking individuals that refuse to receive food from them. In addition, dominants sometimes act aggressively toward individuals that donated food to other subordinate individuals.

Fifth, babblers tend to interfere with others' altruistic behavior. Often, a dominant individual interferes when subordinate individuals are mobbing a predator, serving as sentinels or feeding other individuals (Kalishov, Zahavi and Zahavi 2005; Zahavi and Zahavi 1997). Interference was also reported in helping at the nest as lower-ranking altruists have to "queue" according to their rank in order to help at the nest, and sometimes lower-ranking individuals abort their attempt to help at the nest (or stop helping if they already started to) due to the approach of a dominant individual (Carlisle and Zahavi 1986).

Sixth, the degree of altruistic behavior among babblers seems to be affected by group compositions and the level of competition between potentially breeding individuals. Ridely (2003) found that adult babblers of breeding potential performed more altruistic acts when other adults of breeding potential were present; and that during the breeding season, potential breeders increased the intensity of their altruistic behavior, such as feeding chicks or serving as sentinels. Zahavi and Zahavi (1997: chapter 12) also found that the alpha male tends to mob predators closer, and to interfere in other babblers' mobbing more often, when the alpha male faces more competition from subordinates in his group, in comparison to cases that the alpha male has little competition within his group.

A second line of support to HP comes from studies of sexual selection. The handicap principle has been applied in numerous studies of sexual-selection and costly signaling. Johnstone (1995) has surveyed the pertinent studies and concludes

that “A review of the published data reveals some support for the idea of adaptive choice and honest advertisement ... very few studies report a failure to find such correlates of display or any such benefits” (1995: 51). It is interesting to note, however, that altruistic behavior is not one of the costly signaling mechanisms that Johnstone elaborates upon in his extensive paper, although altruistic behavior is a most prominent mate-selection strategy among babblers and other species as well. Still, Johnstone does bring together a wide array of studies that demonstrate that HP is prominent in mate-selection processes.

Finally, the HP has been applied to cases of altruistic behavior among small-scale hunter-gatherer groups, particularly to foraging strategies and food sharing among non-kin (e.g. Smith and Bliege Bird 2000; Bliege Bird, Smith and Bird 2001; Sosis 2000). These studies have also compared HP with alternative theories and found that HP is best suited to explain why high-quality hunters work harder and receive less calories-intake from their catch or receive no calories at all. Furthermore, it was found that successful hunters, and other group members with high social prestige, “enjoy higher age-specific reproductive success than other Meriam men” (Smith and Bliege Bird 2000: 253; compare to the similar findings of Kaplan and Hill 1985a, 1985b; see also Hawkes 1991). These findings are consistent with HP in the sense that unequal food sharing can be seen as a show-off altruistic display that demonstrates ones qualities as a hunter and contributes to the donors’ reproduction.

3. Reciprocal Altruism Theory

The influence of Robert Trivers’ (1971) ‘reciprocal altruism’ theory on biology, and the study of altruism in particular, had been immense. Fifteen years after the publication of Trivers’ seminal work, ‘the evolution of reciprocal altruism’, a workshop was arranged at UCLA designed to celebrate and assess the empirical and theoretical contributions of RA theory. The workshop papers were assembled in a special issue of *Ethology and Sociobiology*. The editors of this collection note that despite the many problems which have accompanied the application and testing of RA, “it seemed clear to the participants that many problems in behavioral ecology have been profoundly influenced by the theory of reciprocal altruism. Since the theory was originally formulated, a great deal of critical thinking has been stimulated ... [and] much of our thinking and the analysis of data would remain less clear were reciprocal altruism not considered” (Taylor and McGuire 1985: 70).

Indeed, the impact of RA theory on the study of altruism cannot, and should not, be underestimated or ignored given that RA is perhaps the most widely cited theory in the study of altruism and of other phenomena. However, the *contribution* of RA theory, in the sense of advancing our understanding of the puzzle of altruism and enabling one to explain and make useful predictions about the observed phenomenon, is a different matter altogether.

While we acknowledge the importance and originality of RA theory in invigorating the biological thinking about altruism and our understanding of related issues such as cheater detection and game theoretic modeling, we nevertheless believe that this theory does not properly solve the puzzle of altruism, as we explain below.

3.1 The Theoretical Construct

Reciprocal altruism theory proposes that altruistic behavior is possible if, and only if, the individual performing the altruistic act is reciprocated in a later instance by the recipient of the altruistic act. Individuals that behave altruistically toward one another in such a way are supposed to reap long-term benefits that non-altruistic individuals are excluded from. Hence, such a strategy can out-perform more selfish strategies and prevail. RA can evolve given three conditions: “(1) when there are many such altruistic situations in the lifetime of the altruist, (2) when a given altruist repeatedly interacts with the same small set of individuals, and (3) when pairs of altruists are exposed ‘symmetrically’ to altruistic situations, that is, in such a way that the two are able to render roughly equivalent benefits to each other at roughly equivalent costs” (Trivers 1971: 37). Similar restrictions have been proposed by Axelrod and Hamilton (1981) and Axelrod (1984).

Theoretically, these efforts agree that RA (or an equivalent Tit for Tat strategy) cannot evolve from a single altruistic mutation within a group of non-cooperators and it is clear that altruistic behavior cannot invade a group of selfish non-cooperators unless altruists come in clusters (see especially Axelrod 1984). The difficulties of where and how pre-existing clusters of altruistic cooperators come from have been typically downplayed as if it was of lesser theoretical importance. Alternatively, the existence of such pre-existing clusters of altruists was ascribed to a consequence of genetic drift, albeit it can hardly be the case that all the array of altruistic manifestations in various species originated in genetic drift.

3.2 *The (lack of) Empirical Support to Reciprocal Altruism*

Regardless of these basic theoretical problems and restriction, RA can and should be assessed primarily according to its explanatory power and its success in yielding empirical findings that support the theoretical apparatus. A theory that proposes to explain an observed phenomenon may work well in abstract models but can nevertheless fail to yield supporting empirical evidence. Accordingly, in this section we critically examine the empirical evidence that has been reported in support of RA.

Despite the wide acceptance of the theoretical logic of RA and its supposedly explanatory power, it is interesting to note that the theory has yielded very few documented supportive studies (similar points were made by Taylor and McGuire 1985; Koenig 1988; Hammerstein 2003). And, in many instances that were first framed as reciprocity it was found in later inspection that they are actually examples of pseudo-reciprocity such as mutualism, symbiosis, or that those instances only reflect a maximizing strategy within ecological constraints and hence do not involve altruism at all (see also Connor 1986). The few examples that are commonly cited as cases of RA are very problematic as well.

To begin with, the three examples that Trivers (1971) presented in his original paper have all been found to be inaccurate. The case of symbioses between fish and cleaning-fish is not actually a case of reciprocity because there are no real fitness costs to either the fish or the cleaning-fish, both will be worse-off without each other and hence this phenomenon involves no altruism. The second example that Trivers cites is the case of alarm calls in birds. Trivers (1971: 44) proposed that by sounding the alarm call a bird “tends to prevent predators from specializing on the caller’s species and locality.” However, as Koenig (1988: 74) remarks, this example “as proposed by Trivers involves no reciprocal exchange, and thus is certainly not RA” because by making the alarm call the bird reduces its own risks of being eaten and hence benefits directly from the act. The final example proposed by Trivers is reciprocity in humans, to which we shall devote section 4.

Other attempts to confirm RA by empirical studies were equally problematic. For example, Packer’s (1977) study of coalition dynamics among Olive Baboons (*Papio anubis*) is reported to support RA theory because males were observed to come to each other’s help in a seemingly reciprocal manner. This study, however, lends very weak support to RA because the particular costs and benefits from joining and refraining from joining a coalition are unclear, as Packer himself admits.

Moreover, it is not clear from Packer's study whether helping was indeed reciprocal in the symmetrical stipulation proposed by Trivers or whether in actual fact one individual helped his coalition partners more often or less often than others. Packer also did not control the results for the effects of the individuals' rank and status.

Another well cited example that supposedly supports RA is food sharing among bats. Wilkinson (1984; see also Wilkinson 1988 and DeNault and McFarlane 1995) conducted a study of free-ranging vampire bats (*D. rotundus*) and found that out of 110 cases of regurgitation blood donations 21 were certain to be among non-kin (67 other cases were between mother and offspring which are actually cases of parental care, not RA). However, Wilkinson does not present data to determine whether these donations are symmetrical or not, nor does he control for the rank of the individuals. Furthermore, in order to test whether reciprocity may be involved, Wilkinson conducted a separate *lab* experiment where he matched starved bats with satiated bats and found that "bats that received blood later reciprocated the donation" and also that "significantly more often than expected had exchange occurred randomly" (1984: 184).

It is most interesting that this paper is widely cited as an empirical support to RA. First the number of altruistic acts reported among the free-ranging bats is quite small and all it reveals is that 21 cases of blood donations were among non-kin. Second, important conditions of RA have not been tested for, either in captive or free-ranging bats, such as whether donations were symmetrical or not in the sense that some individuals may have donated blood more often than others. Finally, the anomaly of random blood donations which has been reported in the study is very peculiar given that according to RA individuals are supposed to be restrictive as to who they cooperate with and vice versa. Hence, this study bears little support to RA. The study shows that bats tend to donate blood to kin as well as to non-kin, which is precisely what needs explaining, not a proof of any theory.

Other well cited examples of RA are experiments with fish. Milinski (1987) conducted an experiment about patterns of predator-inspection in sticklebacks and found that fish that face a straight mirror are more courageous than those facing an angled mirror, the difference being that the former seemingly simulates a cooperative move and the latter a defecting move. The assumption of the study is that by coming closer to the predator sticklebacks benefit from the information received such as the predator's intentions, distance from the group, etc. Dugatkin

(1988) makes a similar experiment with guppies and reaches similar conclusions based on the same assumptions.

Whether or not the payoffs are indeed the information gathered on the predator, there is an alternative explanation to this type of behavior, which is in line with HP. Predator inspection in fish may be a competition and a show-off display of qualities that pertains to sexual selection. This explanation fits well with the results of these experiments. This is probably the reason sticklebacks confronting a straight mirror spent twice as much time in the front-half of the water-tank in comparison to fish facing the angled mirror, notwithstanding the differences between bolder and more cautious fish (Milinski 1987). The costs entailed are the danger of being eaten or injured, and the qualities being advertised are traits which are relevant to mate selection such as agility in escaping predators and physical condition.

Our interpretation is reinforced by an additional experiment conducted by Dugatkin and Alfieri (1991) in which braver fish were found to be more socially-desired for up to four hours after performing the show-off display in comparison to more timid fish. Whereas Dugatkin and Alfieri interpret this result as a preference for cooperators, it is equally possible to interpret their results as demonstrating preference for mates (if females sought the proximity to the bolder fish) or some kind of hierarchy formation if only males were involved (see also Roberts 1998). It is difficult to determine the reasons why fish seek proximity to the braver fish because the experiment does not clearly specify the sex of the fish involved.

Lastly, reciprocal altruism has been reported in the case of allogrooming in impala (*Aepyceros melampus*). Hart and Hart (1992) report that allogrooming among impala is highly reciprocal in the sense that individuals alternate roles as donor and recipient of bouts of oral grooming to each others' head and neck. Hart and Hart also argue that allogrooming is reciprocal because allogrooming is equal in quantity, and that grooming is unconditioned by genetic relatedness or rank. Nevertheless, this study is most problematic in providing support to RA and in fact contradicts RA in some important respects.

One central condition of RA is a delayed-return of the benefit being donated, whereas in impala the act of giving and receiving is almost simultaneous. Furthermore, it is widely known that allogrooming is not necessarily a 'favor' but may serve as a mechanism to test the social bond (see below).

More seriously, though, are violations of the symmetry condition. Hart and Hart (1991: 1075) report that in certain cases, such as allogrooming among adult

female impala, “it became evident that the allogrooming encounters were highly reciprocal.” However, the evidence that Hart and Hart present hardly support this conclusion and some of their data actually contradict RA assumptions altogether. For instance, in cases of allogrooming in adult female impala, of the four pairs of allogrooming encounters only one pair is truly reciprocal in that the female impalas delivered 42 strokes each. The three other typical encounters were asymmetrical and favored one individual with 9 or 12 bouts more than the other individual received (ibid, 1076). This discrepancy is significant because it constitutes between half to a quarter of what the less-groomed individual received. Hence, this can hardly be considered as reciprocity in quantitative terms.

According to RA, persistence of an asymmetrical behavior confers fitness benefits on one individual and reduces the other animal’s fitness, and hence the more altruistic individual will be selected-out. This issue cannot be further explored because Hart and Hart intentionally sought to “avoid sampling the same animal more than once, except when one subject allogroomed consecutively with different partners” (1992: 1074). Therefore the study does not reveal if some impalas are persistently more altruistic than others in their grooming behavior, but for this same methodological reason the study cannot support RA.

3.2 Reciprocity, Indirect Reciprocity and the Handicap Principle

Another study that was reported to support the theory of RA was conducted on the grooming and alliances relationship among Vervet Monkeys. Seyfarth and Cheney (1984: 542) argue that the “function of grooming among non-kin is to increase the responsiveness of others to the groomer’s subsequent solicitation for aid.” This study proposes a different type of reciprocity in which the benefit paid back is in a different coin in comparison to the type of benefit received. Similar reasoning has been suggested by Frans de Waal (2003).

However, it is important to emphasize that this is no longer the theory of RA as Trivers proposed it in 1971 but rather a different type of logic or even a new theory which can be termed as ‘indirect reciprocity’ (Nowak and Sigmund 1998) or as ‘service economy’ (de Waal 2003). In indirect reciprocity, the costs and benefits to the altruist and to the recipient become ambiguous, they need not be of the same type or degree, and altruistic acts can be reciprocated by the donor’s group and not necessarily by the recipient.

There are several problems with stretching the definition of reciprocity in this way. First, the act of reciprocating an altruistic act can take on any possible

form of benefit which is received (directly or indirectly) from any individual. Hence the term 'reciprocity' is trivialized and is made synonymous with the term 'beneficial' (see also Koenig 1988; and Connor 1986). Indirect RA becomes practically a vacuous theory and almost devoid of predictive power. The reciprocal benefits that are supposedly repaid to the altruist can range from social recognition to abstention from taking an action against the donor. As Koenig (1988: 76) aptly puts the problem with regard to the latter, "the list of behaviors in which individual do *not* engage at any one time is a long one, and limited largely to the imagination of the investigator." Indirect RA cannot really guide the researcher in looking for concrete benefits received by the altruist, and this factor may become rather arbitrary and explained post-hoc.

In the case of grooming behavior, for example, it may very well be that the correlation between grooming and responsiveness to solicitation for help results from an already strong bond between individuals, and that grooming is not the cause for helping but a *reflection* of the strong bond between individuals or their way to measure and sustain their social bond (see for instance the social role of grooming among chimpanzees in de Waal 1982).

Furthermore, some scholars do not regard grooming as an altruistic act at all. Alternative explanations for grooming behavior are establishing and maintaining social bonds (Carpenter 1942; Stambach and Kummer 1982; Matheson and Bernstein 2000) and tension reduction (Terry 1970; O'brien 1993). Similar alternatives exist with regard to allopreening, which is considered a parallel act of grooming in birds: aggression reduction (Harrison 1965; Zacharias and Mathew 1998; Kober and Gaston 2003); testing the social bond (Zahavi 1977b; Zahavi and Zahavi 1997; Dattner 2005) and as means of communication (Dattner 2005). Hence, the explanation that Vervet monkeys (and also primates) indirectly reciprocate acts of grooming in other aspects of their behavior seems rather problematic as a support of RA theory, and in any case grooming is not necessarily causally related to the benefits identified in Seyfarth's and Cheney's (1984) research.

Finally, it is also worth noting the similarity and difference between the HP and 'indirect reciprocity by image scoring' (Nowak and Sigmund 1998), namely explaining the benefits of altruism as pertaining to the altruist's social image as a worthy partner for cooperation. In first appearance, these two explanations may look similar but in fact the logic behind them is substantially different. First, according HP, individual compete over doing altruistic acts and often fiercely refuse to be

“reciprocated.” Second, the social image of the recipient, according to HP, may diminish as a result of receiving help. Third, altruistic acts, according to HP, are important not only to attract collaborators but also to deter rivals. Deterrence does not require reciprocity. Accordingly, individuals have a *direct* interest to behave altruistically toward others and not an indirect benefit of creating a social image as a good cooperator to of receiving indirect help from others.

3.3 Conclusions of the RA studies

The review of the empirical studies which supposedly support RA unravels a most disturbing state of affairs regarding the empirical support, or rather lack thereof, of RA theory and its explanatory power. The theory was not convincingly exemplified in any of the well known studies reviewed above. There are no convincing studies that demonstrate paradigmatic reciprocal altruism according to the stipulations presented by Trivers (1971). These studies share similar problems of violating or not validating the most important condition of RA which is *reciprocity* itself. Reciprocity has been defined by Trivers, and should be defined, as *equal* costs and benefits being conferred. If fewer benefits are conferred relative to the benefits received, this amounts to cheating or defection behavior, and hence not confirming the RA hypothesis. If more benefits are being donated relative to the benefits received, this amounts to altruism on part of the individual, and this is precisely the type of phenomenon that is under consideration, not a proof of the RA hypothesis.

Some of the studies reviewed here ignore the distinction between altruism and reciprocity or conflate the two. Consequently, some of these studies interpret an exchange of benefits between non-kin as proof of RA. However, as argued above, to identify an altruistic exchange between non-related individuals is only to identify and restate the phenomenon under investigation in other words. The fact that two individuals are conferring benefits on each other, even of the same type, can also be interpreted as a competition between these two individuals, not only as cooperation.

One (but not the only) way of distinguishing and deciding between the two alternative interpretations is to observe whether the behavior under question is indeed symmetrical or not. As has been argued above, the studies of RA have not validated this crucial factor. A more important challenge for RA is to explain why individuals interfere, inhibit and act aggressively toward others’ altruistic behavior as shown in the Zahavis work (see section 2.1).

Interestingly, none of these studies tested the alternative HP hypothesis. Data was not collected in order to test whether the abovementioned instances of

altruism can be classified as behavior that pertains to sexual-selection or as competitive behavior between individuals, rather than as instances of cooperation. The lack of reference to HP as an alternative explanation of altruistic behavior is, to some extent, also related to the immense influence of RA on biological thinking and the way RA theory framed the collection of data about altruism among non-kin. That is to say, the huge success of RA in convincing the academic community in the correctness of this hypothesis has constrained, and to some degree also impaired, other ways of thinking about altruism.

Problematically, the lack of supportive empirical evidence has not decreased the acceptance of RA. Hammerstein (2003) has pointed out this problem by asking ‘why is reciprocity so rare in social animals.’ His answer to this peculiar state of the discipline is that “Some theoretical ideas appear to be so compelling that the lack of supportive evidence is indulged by major parts of the scientific community. This criticism applies to current thought in evolutionary biology regarding cooperation in repeated interactions” (Hammerstein 2003: 84). Hammerstein’s criticism is heightened by our research.

4. Altruism in Humans and the Case of One-Shot Altruism

Homo sapiens appear to be the species most capable of practicing reciprocity, indeed perhaps the only species capable of genuine reciprocity. Human beings are cooperating in an array of activities in order to achieve mutual goals and interests. The prevalence of this behavior among human beings is what caught Trivers’ attention in offering RA theory: “Naturally, I thought of ‘you scratch my back – I scratch yours’, which is the folk expression for reciprocity” (see interview with Trivers in de Waal 1996: 25). Indeed, Trivers’ main examples of RA in his 1971 paper come from interpretation of human behavior. Evidence of reciprocity among humans is the main evidentiary pillar of RA. Trivers mentions examples such as hunter-gatherers’ practices of cooperation, emotions of friendship, and Prisoner’s Dilemma experiments.

Trivers has assumed that evolution has shaped the human mind with specific psychological mechanisms that are associated with reciprocity such as a complex regulative system that enables one to selfishly exploit some opportunities by cheating: “namely, when the partner will not find out, when he will not discontinue his altruism even if he does find out, or when he is unlikely to survive long enough to reciprocate adequately” (Trivers 1971: 48). Trivers also cites some Prisoner’s Dilemma game experiments to support the idea that reciprocity is not just an

ultimate factor of evolution but also a proximate psychological mechanism which motivates and guides people in their actions.

However, once RA is applied to human affairs, the results are not so consistent with the theory and with Trivers' predictions. For instance, in the case of hunter-gatherers (viz. Kalahari Bushmen) that Trivers mentions, it is actually the case that a key altruistic behavior of meat-sharing among hunter-gatherers cannot be explained according to RA. Food sharing among hunter-gatherers is significantly non-reciprocal in the sense that some hunters are persistently more successful, and provide more food to the group, than others. Hence, successful hunters obtain the least calories-return rate relative to the effort spent in obtaining the game (Bird et al. 2001; Hawkes 1991; Hawkes et al. 2001; Kaplan and Hill 1985a, 1985b; Lee 1979: chapter 8; Wiessner 1996). These results defy RA predictions that meat sharing should be reciprocal. RA can only be salvaged by introducing concepts of *indirect-reciprocity*, which involve some other types of benefits that were not predicted by, or included in, the initial model (compare Nowak and Sigmund 1998; Smith 2004).

The anomalies of RA predictions become even more pronounced when experiments test for the opportunities to exploit and cheat others in repeated and non-repeated Prisoner's Dilemma games. What these experiments demonstrate is that, more often than not, people do cooperate and behave altruistically even if they can maximize their monetary gains by cheating or by defecting without the other player knowing about it or about their identity. Many people choose the altruistic strategy although they find themselves in one-shot interactions when the rational and selfish option is to cheat or defect, and people are willing to endure costs (e.g. electric shocks or monetary costs) even though the people they are helping are not in a position that enables them to reciprocate in the future (Batson 1991, 1992; Frank 1988; Gintis 2000; Gintis et al. 2003). These results are contradictory of the predictions made by Trivers (1971: 48) and game theorists (Axelrod 1984).

Altruistic behavior in one-shot encounters is, indeed, the most difficult aspect to explain in the altruism puzzle. It is difficult to explain altruistic behavior in one-shot encounters according to RA because by the very nature of this encounter the other individual cannot, or is unlikely, to reciprocate: for instance, in cases donors help a person they will not see again, in anonymous altruism, in cases of altruism with no witnesses, or when donors cooperate and depend on others not to exploit them. HP also faces an interesting dilemma: if altruism is a form of communication it

is not clear who the donors are advertising their qualities to, given that the donors' groups are not watching.

One way to explain one-shot altruism in Prisoner's Dilemma experiments and in real life occurrences is to argue that one-shot altruism is a maladaptation. Trivers (1971) argues that reciprocal altruism is a behavior that had developed during the Paleolithic era when *Homo sapiens* and its predecessors lived in small close-knit groups and when instances of one-shot encounters were probably rare. However, continues Trivers, in modern mass societies, and in conditions in which people meet many other people only once, the predisposition for reciprocal altruism might become a maladaptation. In these new conditions, Trivers says, not only is altruism in one-shot encounters a maladaptation it is far worse than that: "Evolution actually is wiping these people out – it just hasn't finished the job yet" (*New Scientist* 2490, March 2005, p. 33).

We find Trivers' solution to one-shot altruism and unconditional cooperation as an unsatisfactory solution to the explanation of this phenomenon. First, there are evidence suggesting that one-shot encounters were prevalent also in the Paleolithic era when members of small nomadic groups used to meet other groups' members very infrequently due to constant migration, and hence situations of one-shot encounters that involved altruistic behavior (e.g. trusting the other individual, trading with members of other groups, etc., see Lee 1979) could have existed in the distant past as well.

Second, "maladaptation" is a term that could easily be applied to explain away many phenomena whose origins and functions are not fully understood. Maladaptation may then be predicted to disappear in the future. However, one-shot altruism and unconditional cooperation are too prevalent among human beings to be tagged as a maladaptation, especially since there is no indication that this behavior is becoming less common. Furthermore, one-time encounters which may involve altruism between individuals in mass sedentary societies have been a common theme for several thousand years (approximately 6000 years before present), and one-time encounters were most probably very common also for at least 100,000 years before sedentary life began, namely during the nomadic life of the Paleolithic era. Hence, it is difficult to argue that one-shot altruism is a new phenomenon which is not adapted to modern conditions and that will eventually disappear. The high frequency of altruism, either in one-shot encounters or towards people who cannot repay the benefit, testifies that one-shot altruism is adaptive.

Lastly, a very high percentage of people do choose the altruistic strategy of cooperation and altruism over selfish strategies in Prisoner's Dilemma games. While it is possible to say that this is a symptom of the abovementioned maladaptation and that people do not understand the rules of the game because such conditions were absent from humans' evolutionary past, this argument seems weak because people do in fact seem to understand the rules of the game well and change strategies if non-cooperators and defectors try to exploit them (Fehr and Gächter 2002; Gintis et al 2001, 2003).⁸

In our view, the logic of altruistic behavior in one-shot encounters is no different than the logic of other altruistic behaviors. Admittedly, altruism towards strangers whom the altruist is not likely to see again may be a behavior which is unobserved by the altruist's group (e.g. when the altruistic individual is traveling abroad), but the group may learn about the altruist's traits in other ways. For example, it is quite common to hear an anecdote from a traveler who informs the listeners about someone's altruistic deeds while traveling in a distant place, although the donor was not compelled to do so and despite the voluntary risks involved (lending money to a tourist who had all his bags stolen, helping a wounded animal, helping the sick in a distant village, etc.).

But beyond personal testimonies, unconditional altruistic behavior toward strangers or in distant places may be beneficial to the altruist given the advanced and sophisticated modern means of communication, and combined with people's strong tendency to reward altruistic individuals by advertising the altruistic act and thanking donors in other (public) ways.

The occasional discovery of one's altruistic behavior toward strangers or in distant places can serve as a very reliable signal about the individual's traits, and as a consequence may bestow him or her with great social recognition and prestige (e.g. exposing an altruistic individual in the public media tends to cause potential partners to seek the altruist's proximity). Similarly, an incidental discovery of one's selfish behavior may ruin one's reputation. In any case, HP helps to explain and to predict altruistic behavior in one-shot encounters, whereas RA and Tit-For-Tat models predict that individuals will behave selfishly and defect (see also Gintis et al. 2003).

⁸ We owe this point to Peter Hammerstein.

To sum up this point, altruism in one-shot encounters is a key in deciphering the altruism puzzle. The ability of the different theories of altruism to explain this aspect of altruism is an important measure to evaluate the accuracy and usefulness of different theories. RA seems to fall short in explaining and in predicting the cooperation and altruism manifested in one-shot encounters, both in Prisoner's Dilemma games and in real life examples. HP, on the other hand, is more consistent with these results and may provide explanations and predictions to one-shot encounters.

5. Conclusion

Human beings are capable of practicing reciprocity just as they are capable of practicing and designing much more complex forms of exchange, communication, technology and knowledge. Signaling-out reciprocity as a discrete *theory* of evolutionary biology is a very interesting idea that has been tested but not yet proven. Trivers' intuition stemming from the folk expression of "you scratch my back – I scratch yours" was shown possible (with adequate restrictions) on mathematical and game-theoretic models.

RA, however, has not fared well in the empirical realm. RA has not been validated and it remains a theory of restricted explanatory and predictive power. Given this state of affairs, it is worth asking what the role of reciprocity is and whether it is indeed a feature that requires a discrete theory. It could very well be the case that reciprocity is part of other complex mental capabilities, such as the ability to buy and sell goods which is hardly reciprocity.

On the other hand, we posit that an explanation to the puzzle of altruism has been around for three decades, albeit this alternative has been mostly unrealized and unrecognized in the biological scholarship. HP has been applied by Zahavi to the puzzle of altruism already in 1977 but in the biological literature HP theory remains restricted to explanations of other costly-signaling phenomena and to mate-selection. We hope that our clarifications of the HP in this paper will contribute to the theoretical and empirical research of altruism.

Bibliography

- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, Robert and D. William Hamilton. 1981. "The Evolution of Cooperation." *Science* 211: 1390-1396.
- Barkow, Jerome H. 1989. *Darwin, Sex, and Status: biological approaches to mind and culture*. Toronto: University of Toronto Press.
- Batson, C. Daniel. 1991. *The Altruism Question: toward a social-psychological answer*. Hillsdale, N.J ; Hove: Erlbaum.
- Batson, Daniel. 1992. "Experimental Tests for the Existence of Altruism." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2.
- Boyd, R. and P. J. Richerson. 1990. "Group Selection among Alternative Evolutionarily Stable Strategies." *Journal of Theoretical Biology* 145(3): 331-342.
- Boyd, R. and P. J. Richerson. 1992. "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups." *Ethology and Sociobiology* 13(3): 171-195.
- Carlisle, Tamsie R. and Amotz Zahavi. 1986. "Helping at the nest, allofeeding and social status in immature arabian babblers." *Behavioral Ecology and Sociobiology* 18: 339-351.
- Carpenter, C. R. 1942. "Sexual behaviour of free- ranging rhesus monkeys *M. mullata*." *Journal of Comparative Psychology* 33: 113-162.
- Connor, Richard C. 1986. "Pseudo-Reciprocity - Investing in Mutualism." *Animal Behaviour* 34: 1562-1566.
- Darwin, Charles. 1859. *On the Origins of Species*. Cambridge, Mass: Harvard University Press.
- Dattner, A. 2005. *Allopreening in the Arabian Babbler*. MSc thesis, Tel Aviv University (in Hebrew, summary in English).
- DeNault, Lisa K. and Donald A. McFarlane. 1995. "Reciprocal altruism between male vampire bats, *Desmodus rotundus*." *Animal Behaviour* 49(3): 855-856.
- Dugatkin, Lee Alan. 1988. "Do Guppies Play Tit for Tat During Predator Inspection Visits?" *Behavioral Ecology and Sociobiology* 23(6): 395-399.
- Dugatkin, Lee Alan and M. Alfieri. 1991. "Guppies and the Tit-for-Tat Strategy - Preference Based on Past Interaction." *Behavioral Ecology and Sociobiology* 28(4): 243-246.
- Fehr, E. and U. Fischbacher. 2003. "The Nature of Human Altruism." *Nature* 425(6960): 785-791.
- Fehr, Ernst and Simon Gächter. 2002. "Altruistic punishment in humans." *Nature* 415(6868): 137-140.
- Frank, Robert H. 1988. *Passions within Reason: the strategic role of the emotions*. New York ; London: Norton.
- Gintis, H., S. Bowles, R. Boyd and E. Fehr. 2003. "Explaining altruistic behavior in humans." *Evolution and Human Behavior* 24(3): 153-172.
- Gintis, H., E. A. Smith and S. Bowles. 2001. "Costly Signaling and Cooperation." *Journal of Theoretical Biology* 213(1): 103-119.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206(2): 169-179.
- Grafen, Alan. 1990. "Biological Signals as Handicaps." *Journal of Theoretical Biology* 144(4): 517-546.
- Grafen, Alan. 1990. "Sexual Selection Unhandicapped by the Fisher Process." *Journal of Theoretical Biology* 144(4): 473-516.

- Grammer, Karl. 1996. "Symptoms of Power: the function and evolution of social status". in *Food and the Status Quest: an interdisciplinary perspective*. Pauline Wilson Wiessner and Wulf Schiefenheovel (eds.), Providence: Berghahn Books: 68-85.
- Hamilton, D. William. 1964. "The Genetical Theory of Social Behavior, I and II." *Journal of Theoretical Biology* 7: 1-52.
- Hammerstein, Peter. 2003. "Why Is Reciprocity So Rare in Social Animals? A Protestant Appeal". in *Genetic and Cultural Evolution of Cooperation*. Peter Hammerstein (ed.): MIT Press: 83-93.
- Harrison, C.J.O. 1965. "Allopreening as agonistic behaviour." *Behaviour* 24: 161-209.
- Hart, B. L. and L. A. Hart. 1992. "Reciprocal Allogrooming in Impala, *Aepyceros-Melampus*." *Animal Behaviour* 44(6): 1073-1083.
- Hawkes, K., J. F. O'Connell and N. G. B. Jones. 2001. "Hadza meat sharing." *Evolution and Human Behavior* 22(2): 113-142.
- Hawkes, Kristen. 1991. "Showing Off: Tests of an Hypothesis About Men's Foraging Goals." *Ethology and Sociobiology* 12(1): 29-54.
- Hold-Cavell, Barbara. 1996. "The Ethological Basis of Status Hierarchies". in *Food and the Status Quest: an interdisciplinary perspective*. Pauline Wilson Wiessner and Wulf Schiefenheovel (eds.), Providence: Berghahn Books: 19-31.
- Johnstone, R.A. 1995. "Sexual selection, honest advertisement and the Handicap principle: reviewing the evidence." *Biological Review* 70: 1-65.
- Kalishov, Amir, Avishag Zahavi and Amotz Zahavi. 2005. "Allofeeding in Arabian Babblers (*Turdoides squamiceps*)."
Journal of Ornithology 146: 141-150.
- Kaplan, Hillard and Kim Hill. 1985. "Food Sharing among Ache Foragers: Test of Explanatory Hypotheses." *Current Anthropology* 26(2): 223-246.
- Kaplan, Hillard and Kim Hill. 1985. "Hunting Ability and Reproductive Success Among Male Ache Foragers: Preliminary Results." *Current Anthropology* 26(1): 131-133.
- Kober, K. & Gaston, A. 2003. Social interactions among breeding Brünnich's Guillemots *Uria lomiva* suggest constrains in relation to offspring vulnerability. *Ibis*, 145, 413-418.
- Koenig, Walter D. 1988. "Reciprocal Altruism in Birds - a Critical-Review." *Ethology and Sociobiology* 9(2-4): 73-84.
- Lee, Richard B. 1979. *The !Kung San: men, women, and work in a foraging society*. Cambridge: Cambridge University Press.
- Matheson, M. D., Bernstein, I. S. 2000. "Grooming, social bonding, and agonistic aiding in rhesus monkeys" *American Journal of Primatology* 51: 177-186.
- Maynard Smith, J. 1964. "Group Selection and Group Selection." *Nature* 201(4924): 1145-1147.
- Maynard Smith, John. 1976. "Group Selection." *Quarterly Review of Biology* 51(2): 277-283.
- Maynard Smith, John. 1985. "Sexual Selection, Handicaps and True Fitness." *Journal of Theoretical Biology* 115: 1-8.
- Maynard Smith, John. 1991. "Honest Signaling: the Philip Sidney Game." *Animal Behaviour* 42: 1034-1035.
- Milinski, Manfred. 1987. "TIT FOR TAT in stickleback and the evolution of cooperation." *Nature* 325: 433-435.
- Nowak, Martin and Karl Sigmund. 1998. "Evolution of Indirect Reciprocity by Image Scoring." *Nature* 393: 373-377.
- O'brien, T. G. 1993. "Allogrooming behaviour among adult female wedge-capped capuchin monkeys." *Animal Behaviour* 46: 499-510.
- Packer, C. 1977. "Reciprocal Altruism in *Papio anubis*." *Nature* 265: 441-443.

- Pusey, A., J. Williams and J. Goodall. 1997. "The Influence of Dominance Rank on the Reproductive Success of Female Chimpanzees." *Science* 277(5327): 828-831.
- Ridley, A. R. 2003. *The causes and consequences of helping behaviour in the cooperatively breeding Arabian babbler*. PhD Thesis, The University of Cambridge.
- Roberts, Gilbert. 1998. "Competitive altruism: from reciprocity to the handicap principle." *Proc. R. Soc. Lond.* 265: 427-431.
- Seyfarth, Robert and Dorothy L. Cheney. 1984. "Grooming, alliances and reciprocal altruism in Vervet monkeys." *Nature* 308: 541-543.
- Smith, Eric Alden. 2004. "Why do good hunters have higher reproductive success?" *Human Nature-an Interdisciplinary Biosocial Perspective* 15(4): 343-364.
- Smith, Eric Alden and Rebecca L. Bliege Bird. 2000. "Turtle hunting and tombstone opening: public generosity as costly signaling." *Evolution and Human Behavior* 21(4): 245-261.
- Sober, Elliott and David Sloan Wilson. 2000. "Summary of 'Unto Others: the evolution and psychology of unselfish behavior'." *Journal of Consciousness Studies* 7(1-2): 185-206.
- Sosis, Richard. 2000. "Costly signaling and torch fishing on Ifaluk atoll." *Evolution and Human Behavior* 21(4): 223-244.
- Stammach, E. & Kummer, H. 1982. Individual contributions to a dyadic interaction: an analysis of baboon grooming. *Animal Behaviour*, 1982, 30, 964- 971.
- Taylor, Charles E. and Michael T. McGuire. 1988. "Reciprocal Altruism: 15 Years Later." *Ethology and Sociobiology* 9(2-4): 67-72.
- Terry, R. L. 1970. "Primate grooming as a tension reduction mechanism". *Journal of Psychology* 76: 129-136.
- Trivers, Robert. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46: 35-57.
- Turke, Paul W. and L. L. Betzig. 1985. "Those Who Can Do: Wealth, Status, and Reproductive Success on Ifaluk." *Ethology and Sociobiology* 6(2): 79-87.
- Waal, Frans B. M. de. 2003. "The Chimpanzee's Service Economy: Evidence for Cognition-Based Reciprocal Exchange". in *Trust and Reciprocity*. Elinor Ostrom and James Walker (eds.), New Haven: Russell Sage Foundation: 128-143.
- Waal, Frans. B. M. de. 1982. (Rev.). *Chimpanzee Politics: power and sex among apes*. Baltimore; London: Johns Hopkins University Press.
- Waal, Frans. B. M. de. 1996. (7th). *Good Natured: the origins of right and wrong in humans and other animals*. Cambridge, Mass ; London: Harvard University Press.
- Wiessner, Polly. 1996. "Leveling the Hunter: Constraints on the Status Quest in Foraging Societies". in *Food and the Status Quest: an interdisciplinary perspective*. Pauline Wilson Wiessner (ed.), Providence: Berghahn Books: 171-191.
- Wilkinson, Gerald S. 1984. "Reciprocal food sharing in the Vampire bat." *Nature* 308: 181-184.
- Wilkinson, Gerald S. 1988. "Reciprocal Altruism in Bats and Other Mammals." *Ethology and Sociobiology* 9: 85-100.
- Williams, George C. 1966. *Adaptation and natural selection : a critique of some current evolutionary thought*. Princeton: Princeton University Press.
- Wilson, David Sloan. 1977. "Structured Demes and Evolution of Group-Advantageous Traits." *American Naturalist* 111(977): 157-185.
- Wright, Jonathan. 1999. "Altruism as a signal - Zahavi's alternative to kin selection and reciprocity." *Journal of Avian Biology* 30(1): 108-115.
- Wynne-Edwards, Vero Copner. 1962. *Animal dispersion in relation to social behaviour*. Edinburgh: Oliver & Boyd.

- Zahavi, Amotz. 1975. "Mate Selection - a selection for a handicap." *Journal of Theoretical Biology* 53: 205-214.
- Zahavi, Amotz. 1977a. "Reliability in Communication Systems and the Evolution of Altruism". in *Evolutionary Ecology*. Bernard Stonehouse and Christopher Perrins (eds.), London: Macmillan.
- Zahavi, Amotz. 1977b. "The testing of a bond." *Animal Behaviour*, 25, 246-247.
- Zahavi, Amotz. 1995. "Altruism as a Handicap - the Limitations of Kin Selection and Reciprocity." *Journal of Avian Biology* 26(1): 1-3.
- Zahavi, Amotz. 1999. "Babbler altruism and kin selection - a reply to Jon." *Journal of Avian Biology* 30: 115.
- Zahavi, Amotz. 2003. "Indirect selection and individual selection in sociobiology: my personal views on theories of social behaviour." *Animal Behaviour* 65: 859-863.
- Zahavi, Amotz, Avishag Zahavi, Naama Zahavi-Ely and Melvin Patrick Ely. 1997. *The Handicap Principle: a Missing Piece of Darwin's Puzzle*. Oxford: Oxford University Press.
- Zacharias, V. J. & Mathew, D. N. 1998. "Behaviour of the whiteheaded babbler (*Turdoides affinis*)." *Journal of Bombay Natural History Society* 95: 8-14.