The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism

Doron Shultziner and Arnon Dattner

Cogent Biology (2016), 2: 1140857
The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism

Doron Shultziner* and Arnon Dattner

Abstract: This paper reintroduces the Handicap Principle theory as an explanation of altruism. Handicap Principle is a general theory which has been increasingly applied in various disciplines and topics, though strangely it has been missing from both debates and empirical research on altruism in biology. We clarify the logic and principles of HP regarding altruistic behavior based on the costs and benefits to both the donor and the recipient of an altruistic act. We review supportive empirical studies of HP theory. We also compare Handicap Principle theory to Reciprocal Altruism theory and suggest how HP can explain paradigmatic examples of Reciprocal Altruism. Finally, the phenomenon of one-shot altruism is discussed in order to evaluate, and distinguish between, the different explanations of Handicap Principle and Reciprocal Altruism.

1. Introduction
The puzzle of altruistic behavior among human beings and other species has puzzled biologists since Darwin (1859). The puzzle emerged from the need to explain, in evolutionary terms, behaviors that

---

ZOOLOGY | REVIEW ARTICLE

The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism

Doron Shultziner* and Arnon Dattner

Abstract: This paper reintroduces the Handicap Principle theory as an explanation of altruism. Handicap Principle is a general theory which has been increasingly applied in various disciplines and topics, though strangely it has been missing from both debates and empirical research on altruism in biology. We clarify the logic and principles of HP regarding altruistic behavior based on the costs and benefits to both the donor and the recipient of an altruistic act. We review supportive empirical studies of HP theory. We also compare Handicap Principle theory to Reciprocal Altruism theory and suggest how HP can explain paradigmatic examples of Reciprocal Altruism. Finally, the phenomenon of one-shot altruism is discussed in order to evaluate, and distinguish between, the different explanations of Handicap Principle and Reciprocal Altruism.

1. Introduction
The puzzle of altruistic behavior among human beings and other species has puzzled biologists since Darwin (1859). The puzzle emerged from the need to explain, in evolutionary terms, behaviors that

---

Subjects: Evolutionary Psychology; Natural History - Evolution and General Biology; Zoology

Keywords: altruism; handicap principle; costly signaling; reciprocal altruism; indirect-reciprocity; one-shot altruism

1. Introduction
The puzzle of altruistic behavior among human beings and other species has puzzled biologists since Darwin (1859). The puzzle emerged from the need to explain, in evolutionary terms, behaviors that
initially appear to incur fitness costs on the donor while conferring fitness benefits on a recipient (or recipients). Biologists assume that for such a behavior to evolve it must also confer direct or indirect benefits on the donor, otherwise this behavioral tendency would not endure the competitive process of natural selection. Indeed, at the heart of the puzzle of altruistic behavior is the search for an explanation of the logic and benefits in this behavior from the point of view of the individual, whether these benefits are direct or indirect. Several biological theories have been proposed to resolve this puzzle of altruism: Group Selection; Kin Selection (KS, also known as Inclusive Fitness theory); and Reciprocal Altruism (RA). Another theory, which has not received due attention as a theory of altruism, is Handicap Principle. We contend that Handicap Principle theory (HP for short) is a neglected yet very useful theory to explain many instances of altruistic behavior. We argue that it should be taken more seriously and be incorporated into theoretical and empirical studies of altruism.

The contribution of this paper to the study of altruism and cooperation is twofold. First, we highlight key concepts in Zahavi’s (1975, 1977) HP as an explanation of altruistic behavior. We make conceptual clarifications of HP in order to enhance its understanding and applicability to empirical research. We emphasize the underlying competitive nature of altruistic behavior and its potential fitness costs to the recipient, a phenomenon which seldom receives attention in the literature. We demonstrate this competitive dynamic in Arabian babblers and discuss why individuals in general should be recipients at all.

Second, we discuss and compare RA to HP in light of empirical studies, methodological issues, and the phenomenon of one-shot altruism. The motivation for this comparison is to demonstrate how HP has been conceptually and methodologically omitted when it could have provided insight into paradigmatic empirical studies that tested RA and KS explanations alone. We review empirical studies that have been widely cited as supporting RA and reveal common shortcoming in its methodology. We propose how HP could explain certain anomalies and finding in RA studies. We also highlight how RA has been modified into “indirect reciprocity” and why it became a practically unfalsifiable theory. Finally, we discuss altruism among humans and the phenomenon of one-shot altruism. This type of behavior poses interesting explanatory challenges to theories of altruism towards non-kin and to RA in particular.

Due to the known contentious nature of the debate on theories of altruism, it is important to stress that the purpose of this paper is not to develop a new mathematical model of HP. The current models are sufficient (see below). Furthermore, we do not argue that all appearances and examples of altruistic behavior (e.g. a sterile bee defending its hive or foregoing reproduction) need to be explained by HP as instances of altruism or costly signaling. Many appearances of altruism can be explained by simple and straightforward calculations and self-interest of individual-selection, such as parental care (e.g. sacrifice of parents to defend or feed their own offspring, e.g. Trivers, 1974) and as such do not require any theory of altruism. While we elaborate on the elements and characteristics of HP as a general theory, it is worth keeping in mind that we do not argue that all apparent altruism in nature is the result of HP.

2. The handicap principle theory
HP was first presented by Amotz Zahavi in 1975 in relation to mate-selection and was already applied by Zahavi to the question of altruism in 1977. The theory suffered from misunderstanding and a slow start. The situation substantially changed with Grafen’s (1990a, 1990b) two influential papers that gave a mathematical expression to HP (see also Harper’s (2006) and Zahavi’s (2003) reflections on the sociology of science behind the rise of the theory). Ever since the publication of Grafen’s papers, Zahavi’s HP has gained increasing recognition and application in different fields of scholarship, often under the name of Costly Signaling (Harper, 2006; Johnstone, 1995).

Nevertheless, despite its firm mathematical (Grafen, 1990a, 1990b) and game theoretic models (Pilot, 2005; Smith, 1991), HP has not yet been widely applied nor discussed as an underlying principle of altruism although Zahavi (1977) did make that connection explicit (see also Roberts, 1998). Despite its great promise and explanatory robustness, HP has lacked a strong and tenacious cadre
of advocates in the academic discourse on altruism, unlike other theories of altruism. Even those who do mention HP in the current debates on theories of altruism (a rare occurrence) often do not understand the theory, such as Bshary and Bergmüller (2008) who severely misconstrue HP as indirect positive pseudo-reciprocity without realizing that the logic of HP is opposite to reciprocity. In order to show the relevance of HP to altruism we begin with conceptual clarifications of this theory.

2.1. The theoretical construct
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 fitness, not necessarily to its group or kin. The emphasis of the individual’s fitness addresses a common alternative approach that altruistic behavior is “seldom selectively advantageous within groups” and must necessarily be at the between-group level1 (Wilson & Wilson, 2007, p. 338) or kin level2 (West et al., 2007a, 2007b, 2008). According to HP, the benefits of altruistic behavior directly affect the individual’s reproduction; they are not indirect selection benefits (e.g. inclusive fitness or reciprocity). Benefits to kin and the group are possible and quite common outcomes of many instances of altruistic behavior, but not the prime causes of the individual’s behavior. HP has three intertwined assumptions that constitute its theoretical apparatus.

The first initial assumption of HP pertains to phenotypic differences among individuals. These phenotypic differences are primarily (but not exclusively) the result of genetic variability within the group. Due to genetic variability (caused by meiosis and mutation) there are no two individuals with an identical genotype (with the exception of identical twins). Different genotypes give rise to different phenotypes and these differences can become more pronounced due to environmental effects (e.g. differential nutrition or parental care, random causes like injuries, etc.). This assumption essentially means that individuals always differ (at least slightly) in their qualities and traits at different points of their lifespan. This assumption has been a cornerstone in the theory of evolution since Darwin and is found in most biological research.

The second assumption of HP is that many instances of altruistic behavior involve a costly act of communication. Individuals may look morphologically similar but they differ with regards to their qualities. The information about an individual’s qualities is important to potential mates, partners, and rivals. Females and males alike will prefer high-quality partners because this choice would increase their own reproduction. Hence, individuals often benefit from advertising their qualities in order to attract high-quality partners. Individuals competing 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 the same information. Altruism is one such way to communicate with potential mates, partners, and rivals.

Communication of an individual’s qualities is expected to be reliable in order for this system to be functional. A mechanism for communicating qualities may become dysfunctional if it stops to 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 can lose their meaning and may be selected-out (Zahavi & Zahavi, 1997, chapter 4). Therefore, HP proposes that reliable communication involves some sort of costs, costly signals, or handicaps: “a signal is reliable when the difficulty of its performance is related to its meaning in quantity and quality” (Zahavi, 1977a, p. 254). HP proposes a logical connection between the costly signal and its meaning (Zahavi, 1977a; Zahavi & Zahavi, 1997; see also Bliege Bird & Smith, 2005). Namely, costly signals are not random; they tend to convey a message particular to the trait or traits being advertised. For example, individuals that handicap themselves by donating food or nesting others’ eggs signal that they are good foragers and potential reliable partners. Individuals that guard the group more than others advertise their agility and good physical condition. Individuals who give away free gifts signal their wealth and generosity.
HP applies to other forms of costly signaling which are not altruism. For instance, individuals who engage in costly physical displays signal their strength to potential rivals. Individuals who throw away food or valuables in public displays (e.g. in some Indian potlatch ceremonies) communicate qualities such as wealth and power to rivals. Individuals who engage in costly consumption (e.g. show-off huge houses, highly expensive cars) signal their wealth. These are not examples of altruism but of the costly signaling phenomenon more broadly (see Harper, 2006; Johnstone, 1995).

In this context, HP has been identified in an array of other ethnographic behaviors. A useful review has been conducted by Bliege Bird and Smith (2005). In the domain of artistic decoration, Bowser (2000) found that pottery making among Achuar–Quichua women in Ecuador is a form of costly signaling designed to attract high prestige mates for marriage and for maintaining social prestige and political alliances. Wiessner (1984) reported similar findings regarding beaded headbands art by Kalahari women. In the domain of religious ritual and commitment, Boone, Irons, Sosis, and their colleagues have identified many forms of handicaps ranging from time, resources, and various physical handicaps and risks (see Bliege Bird & Smith, 2005, p. 231), all of which were best explained as forms of costly signaling in order to advertise one’s qualities and degree of commitment. Furthermore, HP logic has also been applied to explain monumental architecture in Mesoamerica and embodied handicaps in different cultures (Ibid., pp. 231–232); as well as to explain how moral commitments can be assured (Frank, 1988). Since these are not examples of altruism but of the costly signaling, we do not expand on them in this paper.

The third assumption involves differential costs of signaling. Given the first assumption of phenotypic variability, and supplemented with the second assumption that honest communication of qualities involves costs, a third assumption logically follows that costs of signaling are different to different signalers. Reliable communication is enabled and ensured precisely because the costs of honest signaling of qualities differ between individuals. That is, when differential costs of signaling exist, those individuals who pay the least cost for performing altruistic acts will have the strongest incentive to signal this advantage. From an economic perspective, differential costs of signaling allow individuals to reduce the costs involved in mate selection, competition, and other forms of social exchange (Alvard, 2005, p. 238). This logic works in the following three related ways.

First, costly signaling or self-handicapping (such as altruistic behavior) is less costly to an individual with a better phenotypic quality compared to an individual with a lesser quality (e.g. weaker or less agile). It is less costly for an individual of better quality to bear the same type or degree of altruistic behavior than an individual of lesser phenotypic 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 than more successful individuals to find food and consequently more difficult to bear the costs of altruistic acts such as sharing it or feeding other group members.

Second, an individual of better phenotypic quality can perform higher than average costly signaling relative to other individuals. That is, not only do similar costs differ with regard to their impact on an individual’s survival, but the degree or amount of costly behavior that individuals can endure varies as well. An individual with a better phenotypic quality can perform types or degrees of altruistic behavior (and costly behavior in general) that individuals of different phenotypic quality cannot reach. The ability to endure certain types or degrees of handicaps signals out those individuals that are more fit than those who are less fit at a particular point. This ensures that individuals self handicap only to the degree to which they can cope, i.e. a handicap that honestly reveals their quality. The system is immune to cheating (and unconditioned by social punishment) because individuals that self handicap beyond their real qualities will bear higher costs relative to the higher quality individuals and will lose in fitness. Hence, handicaps increase fitness for honest individuals and will result in a decrease of fitness for cheaters.
Third, altruistic behavior not only increases one's reproduction chances, it can also negatively affect the recipient's (e.g. rival's) fitness, especially in cases of intra-sexual competition. That is, altruism is often used as a competitive strategy that involves costs to 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 decrease because an act of receiving signals and emphasizes the recipient's weakness and inferiority in comparison to the donor. We elaborate on this important aspect of HP in the section below.

2.2. Benefits and costs of altruistic behavior to donor and recipient

The usefulness of HP as an explanation of altruism is also apparent in relation to the costs and benefits to an altruist's and the recipient's fitness. Fitness has two main components: 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 determined and measured by the number of reproducing offspring that the individual has, and not its lifespan (e.g. survival time).

In fact, it is common in nature for many species to employ a strategy of enduring extreme costs to their survival in exchange for reproductive success. Most or all parental-care cases are examples of survival costs exchanged with propagation benefits (e.g. Trivers, 1974). Altruistic behavior involves the same logic: the costs to the donor (the altruist) are primarily in terms of survival risks (costs such as investing energy and time in others, food sharing, risk taking) and the benefits are in terms of higher reproductive chances. Put simply, costs to survival are compensated by gains in reproduction. We should stress that we are talking about cases where there is some sort of cost (minor or major) to the performer of the altruistic act in terms of its time, energy, resources, risk, etc. If no costs are involved, then this is not a case of altruism and no theory of altruism is needed to explain this behavior.

In this context, altruistic behavior is a form of investment (costly communication) in advertising the individual's qualities. Donors benefit directly from sending reliable information about their qualities because this information increases their social prestige, rank, or status in their groups and hence serves to attract mates and allies, and to deter competitors and rivals (see also Bliege Bird & Smith, 2005; Zahavi, 1995). Specifically, it is well known that social prestige and high rank and status within the group are key factors in reproduction. Individuals of higher rank tend to reproduce more and their offspring tend to have higher survivorship rates (Barkow, 1989; Bliege Bird & Smith, 2005; Boone, 1998; de Waal, 1982; Grammer, 1996; Hold-Cavell, 1996; Pusey, Williams, & Goodall, 1997; Smith, 2004; Smith, Bliege Bird, & Bird, 2003; Sosis, 2000; Turke & 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 reproduction (see also Zahavi, 2003).

The costs of this investment are lower to an individual of higher quality than to one of lower quality and this provides the incentive for altruistic behavior. This incentive is independent of group benefits and reciprocity, and it is immune from the free rider problem. The altruist gains a benefit, even if it is not immediate, from the act of displaying. The benefit of the altruist is normally in the form of increased gains in reproduction either directly by way of access to more or better partners or indirectly by way of maintaining or improving the individual's standing within the group (e.g. maintaining allies, attracting more potential partners, and deterring competitors).

On top of the reproduction benefits to the donor, as we indicated above, there are costs to being the recipient of an altruistic act. The recipient may gain the benefits from the donor but the former's chances for reproduction can decrease. In fact, the reproduction benefits of the donor are often at direct expense of the recipient, especially when the interaction involves individuals in a competitive setting over reproduction within the group (see more below). Alternative theories tend to focus on the material and energy dimensions of altruism and neglect the reproduction aspects. Yet, the most altruistic individuals are often the ones with the best reproduction record. That is, the costs of being the recipient of an altruistic act are the ones that matter most to fitness. For this reason, individuals often refuse to be recipients of altruistic acts and may even aggressively react to donors, as will be discussed below.
2.3. Empirical work about altruism and the handicap principle

Zahavi’s work with Arabian babblers (*Turdoides squamiceps*, see Figure 1), a group-breeding songbird, demonstrates several key issues in the nature of altruism. Amotz Zahavi with his wife Avishag Zahavi (see Figure 2), and several other students who have worked with them over more than three decades at the Sheizaf 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 of Arabian babblers are the main (though not only) body of research about HP and altruism. Their studies have shown that babblers compete over performing altruistic acts in several respects. These studies are supported by additional studies of costly signaling and mate choice that are reviewed in this section.

The Zahavis demonstrated that higher ranking males tend to perform substantially more altruistic acts than lower ranking individuals do. Similarly, de Waal (1989) found that top-ranking chimpanzees are amongst the most generous in food sharing. Moreover, high-ranking Arabian babblers 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 (Dattner, Zahavi, & Zahavi, 2015; Zahavi & Zahavi, 1997). Generally, the higher the rank is, the more such altruistic acts that the babblers perform.

Figure 1. Arabian babblers—the only cooperative-breeding bird in Israel.

Source: Arnon Dattner.

Figure 2. Prof. Zahavi and Prof. Avishag Zahavi and the babblers—over 40 years of research.

Source: Arnon Dattner.
The Zahavis also demonstrated that altruistic behavior is almost exclusively unidirectional from a dominant individual to a lower ranking individual. 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 or to close kin (Dattner et al., 2015; Zahavi, 1999). Rather, altruistic acts are directed towards individuals that are closest to the donor in rank (see Alpha male babbler feeding Beta male in subordinated posture, Figure 3). This pattern of unidirectional 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 (Dattner et al., 2015; Kalishov, Zahavi, & Zahavi, 2005).

Furthermore, individuals often refuse to be recipients of altruistic acts, such as being replaced as a sentinel (see high-ranking babbler replacing lower ranking member in Figure 4), being helped at the nest and being fed (Carlisle & Zahavi, 1986; Dattner et al., 2015; Kalishov et al., 2005; Zahavi & Zahavi, 1997). Refusals to be 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. 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 more in Dattner et al., 2015; Kalishov et al., 2005).
More interestingly, babblers tend to be aggressive in response to altruistic acts (Carlisle & Zahavi, 1986; Kalishov et al., 2005). High-ranking babblers tend to be violent toward individuals of lower rank that attempt to feed them. In fact, in certain situations, dominant individuals aggressively respond to a lower ranking individual that tries to feed him or her, even snatching food from the mouth of subordinates and coercing them to become recipients by forcing the food into the subordinate's throat. Alpha males are especially sensitive to such attempts. In the rare occasions 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 (Dattner et al., 2015).

Babblers also 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 et al., 2005; Zahavi & 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 & Zahavi, 1986). The competitive nature of altruism illustrates the fact that altruism is often a mechanism that replaces more dangerous means of maintaining status and communicating one’s qualities such as in physical and violent fights (which also occurs in Arabian babblers, see an injured babbler in Figure 5). The competitive nature of altruism does not mean that all recipients are harmed by the donation. Young offspring who are in need of care obviously benefit from feeding and nesting by non-direct kin, group members benefit from a sentinel’s alarm calls, and a weak or injured adult individual may benefit from help. Yet the competitive nature is quite consistent and clear across all forms of altruistic behavior in babblers (e.g. see a young babbler mobbing a deadly snake in Figure 6).

The degree of altruistic behavior seems to be affected by group compositions and the level of competition between potentially breeding individuals. Ridley (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 social contexts in which the alpha male has little competition within the group.
HP has been also applied to cases of altruistic behavior among small-scale hunter-gatherer groups, particularly to foraging strategies and food sharing (e.g. Bliege Bird & Smith, 2005; Bliege Bird, Smith, & Bird, 2001; Smith, 2004; Smith & Bird, 2000; Smith et al., 2003; Sosis, 2000). These studies also enjoy the advantage of comparing HP with alternative theories, especially reciprocal altruism. The studies found that high-quality hunters work harder and receive less calorie intake from their catch, if they receive any calories from the catch at all. That is, food sharing is typically unconditional. Conditional donations are discouraged and in certain instances are considered culturally inappropriate (Bliege Bird & Smith, 2005, p. 227). These studies found that HP best explains these phenomena due to the direct reputation and reproduction benefits that successful hunters achieve from their altruistic behavior. Successful hunters, and other group members with high social prestige, “enjoy higher age-specific reproductive success than other Meriam men” (Smith & Bird, 2000, p. 253; see also Smith, 2004; Smith et al., 2003). Similar findings regarding subsistence activity have been reported by Kaplan and Hill (1985a, 1985b) and Hawkes (1991). These findings strongly suggest that food sharing, at least among humans, is a show-off altruistic display that signals one’s qualities as a hunter and contributes to the donor’s reproduction.

This wealth of anthropological and multi-disciplinary research is supplemented by studies of mate choice which extends beyond the scope of this paper. Indeed, HP has been applied in numerous studies among many species where close connections were found between handicapped behavior and signaling qualities in mate choice (see review in Johnstone, 1995; see also Ey, Pfefferle, & Fischer, 2007, p. 264).

3. Reciprocal altruism theory

Trivers (1971) reciprocal altruism (RA) theory has had an immense influence on biology and the study of altruism in particular. 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. The workshop papers were assembled in a special issue of *Ethology and Sociobiology*. The editors of this collection note that despite 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 & McGuire, 1988, p. 70)

Indeed, the impact of RA theory on the study of altruism should not be underestimated given that RA is perhaps the most widely cited theory in the study of altruism. RA theory was important and original in invigorating the biological thinking about altruism and our understanding of related...
issues such as cheater detection and game theoretic modeling. In the subsections below we point to several theoretical and empirical problems of RA that could be explained according to HP principles.

3.1. The theoretical construct
Robert Trivers presented Reciprocal Altruism in 1971. His theory proposes that altruistic behavior is possible if, and only if, the individual performing the altruistic act is reciprocated at some point 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 outperform more selfish strategies and prevail. A similar line of reasoning has been proposed in game theoretical models 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 (either genetic or a learned innovation) within a group of non-cooperators and it is clear that altruistic behavior cannot invade a group of selfish non-cooperators unless altruists enter in clusters (see especially Axelrod, 1984). The difficulties of where and how pre-existing clusters of altruistic cooperators come from have been typically downplayed. Alternatively, the existence of such pre-existing clusters or groups of altruists was ascribed to a consequence of genetic drift, although it can hardly be the case that all the array of altruistic manifestations in various species originated in this way (Fehr & Fischbacher, 2003). The resurgence of Group Selection theory is also attributed to, and indeed explicitly builds upon, these theoretic shortcomings (see Wilson & Wilson, 2007, pp. 334–336).

When Trivers originally proposed RA he explained that altruism 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, p. 37)

Trivers’ original model and conditions can be referred to as “testable reciprocal altruism.” These straightforward conditions of RA are both logical and most importantly empirically testable or falsifiable: pairs or altruists are expected to deliver roughly symmetrical costs and benefits to each other in some time lag intervals. Trivers clarified that the “most important parameters to specify for individuals are how many altruistic situations occur and how symmetrical they are; and these are the most difficult to specify in advance” (1971, p. 38). Trivers also emphasizes the condition of symmetry between two individuals in addressing the challenge of cheating. For example, in his discussion about “subtle cheating” in altruism he writes that it also “involves reciprocating, but always attempting to give less than one was given, or more precisely, to give less than partner would give if the situation were reversed” (1971, p. 46). Trivers notes that it is complicated to detect altruistic imbalances and to measure them over a lifespan of individuals, but the thrust of the argument is that symmetry is important between two altruists. We use the original model and its conditions as a yardstick to evaluate paradigmatic empirical studies of RA and how HP may be relevant as an explanation of them (Section 3.2). We also discuss deviations and developments of RA that have made the model practically unfalsifiable (Section 3.3).

3.2. Empirical support of reciprocal altruism?
RA can be assessed primarily according to its explanatory power and its success in yielding empirical findings that support the theoretical apparatus. In the course of this research, we reexamined the empirical evidence that has been reported in support of RA. We found critical methodological issues that have not yet been addressed in the literature and which also shed light on the possible usefulness of HP in explaining altruism.
Despite the wide acceptance of the theoretical logic of RA and its supposed 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 (1988), Koenig (1988), Hammerstein (2003), and West et al. (2007a). Additional research on many reported instances of reciprocity found that they were actually cases of pseudo-reciprocity such as mutualism, symbiosis, or that those instances only reflect a maximizing strategy within ecological constraints and hence do not involve costs at all (Connor, 1986). Despite little supportive empirical evidence, RA remains widely accepted. Hammerstein (2003) has pointed this out by asking “why is reciprocity so rare in social animals” and answering that “[s]ome theoretical ideas appear to be so compelling that the lack of supportive evidence is indulged by major parts of the scientific community” (Hammerstein, 2003, p. 84). Hammerstein’s point is strengthened by our findings.

The examples that are commonly cited as cases of RA are very problematic. For instance, the three examples that Trivers (1971) presented in his original paper have been found to be inaccurate. The case of symbioses between fish and cleaning-fish were not confirmed as reciprocity because there are no real fitness costs to either the fish or the cleaning-fish, both will be worse-off (or die) without each other and hence this phenomenon involves no altruism on their side (Connor, 1986; Koenig, 1988). The second example that Trivers cites is the case of alarm calls in birds. Trivers (1971, p. 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, p. 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. In fact, even if making the alarm call does put the sentinel at risk, as in Arabian babblers, the distribution of risk is far from symmetrical and the alpha male performs it far more than other group members and some individuals are prevented and hindered from performing it (see Dattner et al., 2015). The final example proposed by Trivers is reciprocity in humans, which we shall discuss in Section 4.

Other attempts to confirm RA by empirical studies were equally problematic. We surveyed the literature looking for cited examples in support of RA. Packer (1977) studied coalition dynamics among olive baboons (Papio anubis) and argued that baboons come to each other’s help in a seemingly reciprocal manner. Another well cited example is food sharing among bats (Denault & McFarlane, 1995; Wilkinson, 1984, 1988). Allogrooming in impala (Aepyceros melampus) has also often been cited as a case of RA (Hart & Hart, 1992).

A closer look into these paradigmatic examples reveals severe problems. None of them established the original reciprocal symmetry stipulation of RA. All three studies contend that RA is established because altruistic behavior was unconditioned by genetic relatedness (namely it could not be explained by KS). Specifically, Packer (1977) and Wilkinson (1984) do not present any evidence to determine the extent of costs, let alone whether they are symmetrical or not, nor do they control for the rank of the individuals. RA is simply assumed because altruism is not between kin. HP is not even mentioned in any of the studies as a possible explanation of the behavior at hand.

Hart and Hart (1992), moreover, provide evidence that actually demonstrates significant asymmetrical donor–recipient ratios. For example, in cases of allogrooming in adult female impala, of the four pairs of allogrooming encounters in the study only one pair is truly reciprocal in that the female impalas delivered 42 strokes each, whereas the three other typical encounters are substantially asymmetrical and favor one individual over the other with 9 or 12 bouts, which constitutes between 25 and 50% of what the less groomed individual received (Hart & Hart, 1992, p. 1076). According to the testable RA conditions, if allogrooming is the benefit being conferred, “pairs of altruists are exposed ‘symmetrically’ to altruistic situations ... in such a way that the two are able to render roughly equivalent benefits to each other at roughly equivalent costs” (Trivers, 1971, p. 37). The reported discrepancy in allogrooming within the impala pairs is not subtle cheating, it is gross cheating. Persistence of such asymmetrical cost–benefit behavior as assumed and reported in Hart and Hart (1992) confers fitness benefits on one individual and reduces the other’s fitness. Conversely,
asymmetrical donations of this sort fit well with HP logic of altruism that the impala pairs are not reciprocating but rather competing.

Another kind of RA research on patterns of predator-inspection in sticklebacks (Milinski, 1987) and guppies (Dugatkin, 1988) maintains that cooperation is conditioned on reciprocity. Yet, the evidence reported in these experiments seems to fit much better with HP than with RA. Sticklebacks confronting a straight mirror (hypothesized in the studies as simulating cooperation) spent twice as much time in the front-half of the water tank compared to fish facing the angled mirror (hypothesized as simulating defection), 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 choice such as agility in escaping predators and physical condition. Hence, predator inspection in fish is probably a competition and a show-off display of qualities that pertains to mate choice, rather than a mechanism required for cooperation in predator inspection. Indeed, when the experiment simulated conditions of competition, the fish behaved more bravely, and in conditions that did not simulate competition, brave behavior was muted. This observed behavior fits far better with HP than with RA.

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 more plausible 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 specify the sex of the fish involved. The fact that in “many fish species one or a few individuals break away from a school and approach a predator to assess the potential danger” (Dugatkin, 1997, p. 357) implies that this behavior is most probably a display of quality or a type of altruistic act if benefits to other individuals are indeed involved. In either case, this behavior directly increases the reproductive chances of bolder and higher quality fish relative to more cautious fish. The examples above illustrate the different data, research design, and interpretations of the same behavior by RA and HP, though the latter has been excluded.

3.3. Reciprocity, indirect reciprocity, and the handicap principle

Since evidence of RA has been scarce, RA scholars found it necessary to loosen the original RA conditions that were noted above and deviate from the key conditions of reciprocal symmetry, or “roughly equivalent benefits ... at roughly equivalent costs” as Trivers (1971, p. 37) first put it. Common examples of this shift are studies on grooming and alliance relationships among vervet monkeys and chimpanzees. Seyfarth and Cheney (1984, p. 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.” Similar reasoning has been suggested by de Waal (2003) in studies of chimpanzees. These studies propose two different types of reciprocity. In the first type, the benefit paid back is in a very different coin or form than that of the benefit received and not even by an altruistic act. In the second type, a different individual than the one that received the benefit pays the benefit back.

In our view this is no longer a testable RA theory but a different type of logic and theory which was termed “indirect reciprocity” (Nowak & Sigmund, 1998) and “service economy” (de Waal, 2003). In this model, the costs and benefits to the altruist and to the recipient are ambiguous, they need not be of the same type or degree, and altruistic acts can be reciprocated by non-altruistic acts (e.g. sex, passivity, and deterrence) and by the donor’s group and not even necessarily by the recipient. There are several problems with this significant and crucial deviation from the testable RA model. By stretching the definition of reciprocity to such an extent, the act of reciprocating an altruistic act can take on any possible form of non-altruistic 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 Connor, 1986; Koenig, 1988). Indirect reciprocity becomes a vacuous theory, devoid of actual predictive power. The reciprocal benefits that are supposedly repaid to the altruist can range from sex with an individual that was not the recipient, group social recognition, to abstention from taking an action against the donor. As Koenig (1988, p. 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 explanations might become post hoc. Most seriously, the deviation from the defining condition of testable RA—roughly symmetrical costs and benefits between the same two individuals that engage in altruistic situations—has yielded an unfalsifiable model. It is arguably impossible not to find some benefit—real or imagined—to the performer of altruistic acts from either the recipient of the act (e.g. passivity or reduced aggression) or unaffected individuals (e.g. being sought as a mate). Without prior theory and conditions about the costs, benefits, and the mechanisms of reciprocity, it is hard to see how this indirect reciprocity model stands to the basic scientific criterion of being falsifiable.

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 existing bond between individuals. Grooming, then, is not the cause for helping but a reflection of this social bond or a way to measure and sustain a social bond (see for instance the social role of grooming among chimpanzees; 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; Matheson & Bernstein, 2000; Stammbach & Kummer, 1982) and tension reduction (O’Brien, 1993; Terry, 1970). Similar alternatives exist with regard to allopreening, which is considered a parallel act of grooming in birds; aggression reduction (Harrison, 1965; Kober & Gaston, 2003; Zacharias & Mathew, 1998); testing the social bond (Dattner, 2005; Zahavi, 1977b; Zahavi & Zahavi, 1997); and as means of communication (Dattner, 2005). Hence, the explanation that vervet monkeys (and other primates) indirectly reciprocate acts of grooming in other aspects of their behavior seems arbitrary and impossible to refute. In short, grooming is not necessarily causally related to the benefits identified in Seyfarth and Cheney’s (1984) and it is problematic as support to RA theory.

Nowak and Sigmund (1998) proposed “indirect reciprocity by image scoring,” that is, explaining the benefits of altruism as pertaining to the altruist’s social image as a worthy partner for cooperation. It is worth noting the similarity and difference between indirect reciprocity and HP. At first glance, these two explanations may look similar but in fact the logic behind them is substantially different. First, according to HP, individuals compete over performing 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 any form of reciprocity by the recipient; yet a deterred individual might accidently be interpreted as indirectly reciprocating by behaving passively or by cooperating with the alpha male or female. According to HP, individuals have a direct interest to behave altruistically toward others and not an indirect benefit of creating a social image as a good cooperators or of receiving indirect help from others. These differences provide clear guidelines for field research, but none of the studies mentioned above compared RA to HP.

3.4. Conclusions from the RA studies

The review of the empirical studies which supposedly support RA reveals a problematic 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 symmetrical reciprocal-altruism. These studies share similar problems of not validating or even violating the most crucial condition of, and test to RA which is reciprocity itself. Reciprocity has been defined by Trivers, and indeed should be defined, as roughly 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 it does not confirm the RA hypothesis. If more benefits are being donated relative to the benefits received then this begs the question of how asymmetrical costly behavior can be maintained in evolutionary terms. This brings us a full circle back to the original puzzle of altruism.

Furthermore, the fact that the exchange of benefits between non-kin has often been interpreted as proof of RA is highly problematic. The reason for this methodological problem should be seen in the context of the 1970–1990 period in which the dominant theories of altruism were KS and RA. Hence, demonstrating that altruism extends beyond kin was stated as proof to RA, even when the criteria of RA were not in fact verified. Yet, to identify asymmetrical benefits between two non-related individuals is altruism that needs to be explained; i.e. it only identifies and restates altruism in other words and it is not proof of RA or any other theory. According to HP, conferral of asymmetrical benefits is often a competition for status between the two individuals.

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. A more important challenge for RA and KS is to explain why individuals often interfere, inhibit and, act aggressively toward others’ altruistic behavior or even their close relatives’ attempts to perform altruistic acts (see above).

Interestingly, none of these studies tested the alternative HP explanation. Data were not collected in order to test whether the above-mentioned instances of altruism can be classified as behavior that pertains to mate choice or as competitive behavior between individuals, rather than as instances of cooperation or inclusive fitness. 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 in research about altruism among non-kin. That is to say, the rallying behind RA has constrained, and to some degree also impaired, other ways of thinking about altruism, even in cases where the most straightforward explanation involves costly signaling or mate choice.

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 cooperate 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, p. 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, p. 48). Trivers also cites Prisoner’s Dilemma game experiments to support the idea that reciprocity is not just a 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 very 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; and hence successful hunters obtain the least calories return rate relative to the effort spent in obtaining the game (Bliege Bird & Smith, 2005; Bliege Bird et al., 2001; Hawkes, 1991; Hawkes, O’Connell, & Blurton Jones, 2001; Kaplan & Hill, 1985a, 1985b; Lee, 1979, chapter 8; Smith, 2004; Smith & Bird, 2000; Smith et al., 2003; Sosis, 2000; Wiessner, 1996). These results explicitly defy RA predictions that meat sharing should be symmetrical. RA can only be salvaged by introducing indirect-reciprocity, which involves some other types of benefits that were not predicted by, or included in, the initial model (compare Nowak & 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. These experiments demonstrate 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 when they find themselves in one-shot interactions in which the rational and selfish option is to cheat or defect. Furthermore, people are even willing to endure costs (e.g. electric shocks or monetary costs) when the people they are helping are not in a position to reciprocate (e.g. Batson, 1991, 1992; Frank, 1988; Gintis, 2000; Gintis et al., 2003). These results are contradictory to Trivers’s predictions (1971, p. 48) and reciprocity game-theorists (Axelrod, 1984).

Altruistic behavior in one-shot encounters is, indeed, the most difficult aspect to explain in the altruism puzzle. This is because by the very nature of this encounter the other individual cannot, or is unlikely, to reciprocate: for instance, in cases where 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 of one’s qualities then it is not clear who the donors are advertising their qualities to and how it will be beneficial, given that the donors’ groups may not be observing the act.

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. Trivers maintains that 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, 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,” argues Trivers (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, it is quite inaccurate that one-shot encounters were uncommon in the Paleolithic era because in that long era of human existence members of small nomadic groups used to meet individuals from other groups very infrequently due to constant migration and the social flux characterizing small nomadic bands of foragers. Situations of one-shot encounters and altruistic behavior such as trading with members of other groups (Lee, 1979) most probably existed among distant nomadic bands during the bulk of human evolution in the Paleolithic era, as shown in genetic studies of those populations (Behar et al., 2008; Churchill, Formicola, Holliday, Holt, & Schumann, 2006; Mellars, 2006).

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, and there is no indication that this behavior is becoming less common. Furthermore, one-time altruistic encounters in mass sedentary societies have been a common theme for several thousands of years as well as encounters between highly
flexible and distant nomadic bands before mass sedentary life came into existence. 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 and potentially beneficial independent of kin and reciprocity factors.

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 understand the rules of the game well and change strategies (sometimes in costly and spiteful ways) if non-cooperators and defectors try to exploit them (e.g. Fehr & Gächter, 2002; Gintis et al., 2001, 2003).

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 unlikely 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 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 highly reliable signal about the individual’s traits, and as a consequence may gain him or her considerable 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). Conversely, an incidental discovery of one’s clandestine selfish behavior may completely ruin one’s reputation. HP helps to explain and to predict altruistic behavior in such one-shot encounters, whereas RA and Tit-For-Tat models predict that individuals will behave selfishly and defect (see also Gintis et al., 2003).

Finally, 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. Given the few empirical studies of reciprocity and their shortcomings, it is worth asking whether reciprocity is indeed a feature that requires a discrete theory of altruism or perhaps simply part of other complex mental capabilities, such as the ability to buy and sell goods which do not involve a biological theory.

5. Conclusion

The explanatory potential of Handicap Principle has not been fully realized in the biological scholarship. Although it has been around for some time, and has proven useful in empirical research across disciplines, HP remains restricted to explanations of costly signaling phenomena and to mate-selection. We hope that this paper will advance the understanding of HP and will facilitate its inclusion and hypothesis testing in future explanations of altruistic behavior. The clarifications of HP, and its comparison to RA, could contribute to research methodology pertaining to altruism in the following ways.

First, future studies of altruistic behavior need to be explicit about the assumed benefits and costs of altruistic behavior at the individual level of analysis. Only by specifying such costs and benefits will it be possible to test alternative hypotheses and to decide which is most relevant for each
case-specific behavior and species-specific behavior. By following this methodological guideline, it would also be possible to explain behavior by simple individual calculations stemming from maximizing strategies or direct benefits to the performer of that act that involve no actual cost to the individual (i.e. no altruism involved) or by lack of choice that may also involve no altruism (e.g. lack of other reproductive possibilities). This sort of explanation would not need to invoke any theory of altruism.

Second, research should also be sensitive to the possible costs and benefits to the recipient of the altruistic act. Research on altruism in the empirical (field) level can obtain much insight from observing the effects on recipients of altruistic acts in terms of viability on the one hand and reproduction on the other. Rather than assume that the conferred benefit (e.g. food, or replacement at sentinel position, reduced risk) is beneficial to the recipient, research must demonstrate that this is so. This focus on the recipient may shed new light on how some recipients are not entirely benefiting from altruism and raise the question of why individuals agree to be recipients of altruism.

Thirdly and related, research should be sensitive to the possible competitive nature of altruistic acts. What might seem to be reciprocity at first sight could be in fact a mechanism of competition that replaced more costly and dangerous (e.g. physical) means of competition into more benign ways that also have some benefits to the group (but not due to such group-level benefits). We hope this paper will encourage scholars to consider these methodological guidelines in research on altruism, and that more scholars will incorporate HP a possible explanation of altruistic behavior.

Acknowledgements
We thank Avishag and Amotz Zahavi for their support and encouragement. We also thank the anonymous reviewers for their very constructive and helpful comments.

Funding
The authors received no direct funding for this research.

Competing interests
The authors declare no competing interest.

Author details
Doron Shultziner1
E-mail: doronsh@hadassah.ac.il
Aron Dattner2
E-mail: Nonydt@gmail.com
1 Department of Politics & Communications, Hadassah Academic College, Jerusalem, Israel.
2 Jerusalem, Israel.

Citation information
Cite this article as: The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism, Doron Shultziner & Arnon Dattner, Cogent Biology (2016), 2: 1140857.

Cover image
A dominant Arabian Babbler male feeds a lower ranking male: altruism is a competition over in-group rank (Hatzeva Nature Reserve in Israel).
Source: Arnon Dattner.
https://youtu.be/LLJgR0f15ZE.

Notes
1. The literature on Group Selection is not discussed in this paper. For further reading see Wynne-Edwards (1962), Wilson (1977), Boyd and Richerson (1990, 1992), Saber and Wilson (2000), Fehr and Gächter (2002), Gintis (2000), Gintis, Smith, and Bowles (2001), Gintis, Bowles, Boyd, and Fehr (2003), Smith (1966, 1976), Wilson and Wilson (2007), Nowak, Tamita, and Wilson (2010), and Nowak and Allen (2015).
2. The literature on Kin Selection is not discussed in this paper. For further reading see Hamilton (1964), Dawkins (1979), West, Griffin, and Gardner (2007a, 2007b, 2008), Bshary and Bergmüller (2008), Liao, Rong, and Queller (2015), Queller, Rong, and Liao (2015).

References
Alvard, M. (2005). Comments on Bliege Bird and Smith’s signaling theory. Strategic Interaction, and Symbolic Capital, Current Anthropology, 46, 238–239.
Axelrod, R. (1984). The evolution of cooperation. New York, NY: Basic Books.
Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. Science, 211, 1390–1396.
http://dx.doi.org/10.1126/science.7466396
Barkow, J. H. (1989). Doran, sex, and status: Biological approaches to mind and culture. Toronto: University of Toronto Press.
Batson, C. D. (1991). The altruism question: Toward a social-psychological answer. Hove: Erlbaum.
Batson, D. (1991). Experimental tests for the existence of altruism. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 2, 69–78.
Behar, D. M., Richard, V., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., ... Rosset, S. (2008). The dawn of human matrilineal diversity. The American Journal of Human Genetics, 82, 1–11.
Bliege Bird, R., & Smith, E. A. (2005). Signaling theory, strategic interaction, and symbolic capital. Current Anthropology, 46, 221–248.
Bliege Bird, R., Smith, E. A., & Bird, D. W. (2001). The hunting handicap: Costly signaling in human foraging strategies. Behavioral Ecology and Sociobiology, 50, 9–19.
http://dx.doi.org/10.1007/s002650100338
Boone, J. L. (1998). The evolution of magnanimity. Human Nature, 9(1), 1–21.
http://dx.doi.org/10.1007/s12110-998-1009-y
Bowser, B. J. (2008). From pottery to politics: An ethnoarchaeological study of political factionalism, the Ecuadorian Amazon. Journal of Archaeological Method and Theory, 7, 219–248.
http://dx.doi.org/10.1023/A:10265010620824
Wynne-Edwards, V. C. (1962). Animal dispersion in relation to social behavior. Edinburgh: Oliver & Boyd.
Zahavi, A. (1975). Mate selection—A selection for a handicap. Journal of Theoretical Biology, 53, 205–214. http://dx.doi.org/10.1016/0022-5193(75)90111-3
Zahavi, A. (1977a). Reliability in communication systems and the evolution of altruism. In B. Stonehouse & C. Perrins (Eds.), Evolutionary ecology (pp. 253–259). London: Macmillan.
Zahavi, A. (1977b). The testing of a bond. Animal Behaviour, 25, 246–247. http://dx.doi.org/10.1016/0003-3472(77)90089-6
Zahavi, A. (1995). Altruism as a handicap: The limitations of Kin selection and reciprocity. Journal of Avian Biology, 26(1), 1–3. http://dx.doi.org/10.2307/3677205
Zahavi, A. (1999). Babbler altruism and Kin selection: A reply to Jon. Journal of Avian Biology, 30, 115. http://dx.doi.org/10.2307/3677251
Zahavi, A. (2003). Indirect selection and individual selection in sociobiology: My personal views on theories of social behaviour. Animal Behaviour, 65, 859–863. http://dx.doi.org/10.1006/anbe.2003.2109
Zahavi, A., & Zahavi, A. (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.