

Putting the Altruism Back into Altruism: The Evolution of Empathy

Frans B.M. de Waal

Living Links, Yerkes National Primate Research Center, and Psychology Department, Emory University, Atlanta, Georgia 30322; email: dewaal@emory.edu

Annu. Rev. Psychol. 2008. 59:279–300

First published online as a Review in Advance on June 5, 2007

The *Annual Review of Psychology* is online at <http://psych.annualreviews.org>

This article's doi: 10.1146/annurev.psych.59.103006.093625

Copyright © 2008 by Annual Reviews. All rights reserved

0066-4308/08/0203-0279\$20.00

Key Words

perception-action, perspective-taking, prosocial behavior, cooperation

Abstract

Evolutionary theory postulates that altruistic behavior evolved for the return-benefits it bears the performer. For return-benefits to play a motivational role, however, they need to be experienced by the organism. Motivational analyses should restrict themselves, therefore, to the altruistic impulse and its knowable consequences. Empathy is an ideal candidate mechanism to underlie so-called directed altruism, i.e., altruism in response to another's pain, need, or distress. Evidence is accumulating that this mechanism is phylogenetically ancient, probably as old as mammals and birds. Perception of the emotional state of another automatically activates shared representations causing a matching emotional state in the observer. With increasing cognition, state-matching evolved into more complex forms, including concern for the other and perspective-taking. Empathy-induced altruism derives its strength from the emotional stake it offers the self in the other's welfare. The dynamics of the empathy mechanism agree with predictions from kin selection and reciprocal altruism theory.

Contents

INTRODUCTION	280
ORIGIN OF EMPATHY	282
LEVELS OF EMPATHY	282
Emotional Contagion	282
Sympathetic Concern	283
Empathic Perspective-Taking	285
UNDERLYING MECHANISMS	286
Perception Action Mechanism	286
Russian Doll Model	287
FROM EMPATHY TO	
ALTRUISM	288
Emotional Contagion	288
Sympathetic Concern	289
Empathic Perspective-Taking	289
EMPATHY AS EVOLVED	
PROXIMATE MECHANISM OF	
DIRECTED ALTRUISM	291
CONCLUSION	292

Sympathy . . . cannot, in any sense, be regarded as a selfish principle.
Smith (1759, p. 317)

Empathy may be uniquely well suited for bridging the gap between egoism and altruism, since it has the property of transforming another person's misfortune into one's own feeling of distress.
Hoffman (1981a, p. 133)

Altruism (biological definition):

behavior that increases the recipient's fitness at a cost to the performers

Ultimate cause or goal: the benefits an organism or its close kin derive from a behavior, hence the probable reason why the behavior was favored by natural selection

Proximate cause: situation that triggers behavior and the mechanism (psychological, neural, physiological) that enables it

INTRODUCTION

Discussions of altruistic behavior tend to suffer from a lack of distinction between function and motivation. This is due to the contrasting emphasis of biologists and psychologists, with the former focusing on what a particular behavior is good for, and the latter on how it comes about.

Evolutionary explanations are built around the principle that all that natural selection can work with are the effects of behavior—not the motivation behind it. This means there is only one logical starting point for evolutionary accounts, as explained by Trivers (2002, p. 6):

“You begin with the effect of behavior on actors and recipients; you deal with the problem of internal motivation, which is a secondary problem, afterward. . . . [I]f you start with motivation, you have given up the evolutionary analysis at the outset.”

This is a perfectly legitimate strategy that has yielded profound insights into the evolution of altruism (e.g., Dugatkin 2006). Unfortunately, however, these insights have not come with a new terminology: Evolutionary biology persists in using motivational terms. Thus, an action is called “selfish” regardless of whether or not the actor deliberately seeks benefits for itself. Similarly, an action is called “altruistic” if it benefits a recipient at a cost to the actor regardless of whether or not the actor intended to benefit the other. The prototypical altruist is a honeybee that stings an intruder—sacrificing her life to protect the hive—even though her motivation is more likely aggressive than benign. This usage of the terms “selfish” and “altruistic” oftentimes conflicts with their vernacular meaning (Sober & Wilson 1998).

The hijacking of motivational terminology by evolutionary biologists has been unhelpful for communication about motivation *per se*. The way to clear up the confusion is to do what Trivers did when he decided that evolutionary analyses require that effects be considered separate from motivation. Conversely, motivational analyses require us to keep motivation separate from evolutionary considerations. It is not for nothing that biologists hammer on the distinction between ultimate and proximate (Mayr 1961, Tinbergen 1963). The ultimate cause refers to why a behavior evolved over thousands of generations, which depends on its fitness consequences. The proximate cause, on the other hand, refers to the immediate situation that triggers behavior, and the role of learning, physiology, and neural processes—typically the domain of psychologists.

Proximate and ultimate viewpoints do inform each other, yet are not to be conflated. For example, primate cooperation is

promoted by social tolerance. Through its effect on food-sharing, tolerance evens out payoff distributions (de Waal & Davis 2003, Melis et al. 2006). Tolerance likely is a proximate mechanism that evolved to serve the ultimate goal of cooperation, which is to yield benefits for all contributors.

Cooperation and altruistic behavior are thought to have evolved to help family members and those inclined to return the favor (Hamilton 1964, Trivers 1971). Regardless of whether this is the whole explanation or not (see Sober & DS Wilson 1998, EO Wilson 2005), the point is that ultimate accounts stress return-benefits, i.e., positive consequences for the performer and/or its kin. Inasmuch as these benefits may be quite delayed, however, it is unclear what motivational role, if any, they play. This becomes clear if we consider more closely what drives directed altruism, i.e., altruistic behavior aimed at others in need, pain, or distress. There are three ways in which directed altruism may come about:

1. Altruistic impulse. Spontaneous, disinterested helping and caring in reaction to begging or distress signals or the sight of another in pain or need.
2. Learned altruism. Helping as a conditioned response reinforced by positive outcomes for the actor.
3. Intentional altruism. Help based on the prediction of behavioral effects. One prediction could be that the help will be reciprocated, hence that the act will produce a net benefit. Since the actor seeks to benefit itself, we may call this intentionally selfish altruism. The second possibility is help based on an appreciation of how one's own behavior will help the other. Since the actor seeks to benefit the other, we may call this intentionally altruistic altruism.

Some directed altruistic behavior is promoted by built-in rewards, such as the oxytocin release during suckling that may underpin maternal care (Panksepp 1998). Empathy-based altruism may have similar in-

trinsically rewarding qualities in that it offers the actor an emotional stake in the recipient's well-being, i.e., if helping the other ameliorates the helper's internal state (see Empathy as Evolved Proximate Mechanism, below). Extrinsic rewards, on the other hand, are less likely to play a role. By definition, altruism carries an initial cost, and positive consequences occur only after a significant time interval (e.g., the recipient reciprocates) or not at all (e.g., care for dependent kin), making for rather poor learning conditions.

Intentionally selfish altruism would require the actor to explicitly expect others to return the favor. Despite the lack of evidence for such expectations in animals, they are often assumed. The common claim that humans are the only truly altruistic species, since all that animals care about are return-benefits (e.g., Dawkins 1976, Fehr & Fischbacher 2003, Kagan 2000, Silk et al. 2005), misconstrues reciprocity as a motivation. It assumes that animals engage in reciprocal exchange with a full appreciation of how it will ultimately benefit them. Helpful acts for immediate self-gain are indeed common (Dugatkin 1997), but the return-benefits of altruistic behavior typically remain beyond the animal's cognitive horizon, i.e., occur so distantly in time that the organism is unlikely to connect them with the original act. This applies to most reciprocal altruism in the animal kingdom.

Once evolved, behavior often assumes motivational autonomy, i.e., its motivation becomes disconnected from its ultimate goals. A good example is sexual behavior, which arose to serve reproduction. Since animals are, as far as we know, unaware of the link between sex and reproduction, they must be engaging in sex (as do humans much of the time) without progeny in mind. Just as sex cannot be motivated by unforeseen consequences, altruistic behavior cannot be motivated by unforeseen payoffs.

The altruistic impulse is to be taken very seriously, therefore, because even if altruistic behavior were partially learned based on

Directed altruism: helping or comforting behavior directed at an individual in need, pain, or distress

Intentional altruism: the altruist deliberately seeks to benefit either the other (intentionally altruistic altruism) or itself (intentionally selfish altruism)

Empathy-based altruism: help and care born from empathy with another

Empathy: the capacity to (a) be affected by and share the emotional state of another, (b) assess the reasons for the other's state, and (c) identify with the other, adopting his or her perspective. This definition extends beyond what exists in many animals, but the term "empathy" in the present review applies even if only criterion (a) is met

Motivational autonomy: independence of motivation from ultimate goals

Perception-action mechanism (PAM):

automatically and unconsciously activated neural representations of states in the subject similar to those perceived in the object

Emotional contagion:

emotional state-matching of a subject with an object

short-term intrinsic rewards or long-term extrinsic rewards, this by no means rules out the altruistic impulse. In fact, it presupposes this impulse given that a behavior's consequences cannot be learned without spontaneously engaging in it in the first place.

This review seeks to restore the altruism within altruism by exploring the role of empathy in the directed altruism of humans and other animals. Some definitions of empathy stress the sharing of emotions, whereas other definitions stress the capacity to put oneself into the other's "shoes." The latter definitions are so top-down, however, that they disconnect empathy from its possible antecedents. We follow a bottom-up approach instead, adopting the broadest possible definition, including mere emotional sensitivity to others. We first consider the various levels of empathy in animals and the underlying perception-action mechanism (PAM) proposed by Preston & de Waal (2002a). After this, we explore the relation between empathy and altruism.

A major question is whether evolution is likely to have selected empathy as proximate mechanism to generate directed altruism. Does empathy channel altruism in the direction that evolutionary theory would predict? So, even though motivation will be kept temporarily separate from evolutionary considerations, in the end the two will meet. Empathy may be motivationally autonomous, but it still needs to produce—on average and in the long run—evolutionarily advantageous outcomes. The central thesis to be argued here, then, is that empathy evolved in animals as the main proximate mechanism for directed altruism, and that it causes altruism to be dispensed in accordance with predictions from kin selection and reciprocal altruism theory.

ORIGIN OF EMPATHY

Empathy allows one to quickly and automatically relate to the emotional states of others, which is essential for the regulation of social interactions, coordinated activity, and cooperation toward shared goals. Even though

cognition is often critical, it is a secondary development. As noted by Hoffman (1981b, p. 79), "[H]umans must be equipped biologically to function effectively in many social situations without undue reliance on cognitive processes."

The selection pressure to evolve rapid emotional connectedness likely started in the context of parental care long before our species evolved (Eibl-Eibesfeldt 1974 [1971], MacLean 1985). Signaling their state through smiling and crying, human infants urge their caregiver to come into action (Acebo & Thoman 1995, Bowlby 1958). Equivalent mechanisms operate in all animals in which reproduction relies on feeding, cleaning, and warming of the young. Avian or mammalian parents alert to and affected by their offspring's needs likely out-reproduced those who remained indifferent.

Once the empathic capacity existed, it could be applied outside the rearing context and play a role in the wider network of social relationships. The fact that mammals retain distress vocalizations into adulthood hints at the continued survival value of empathy-inducing signals. For example, primates often lick and clean the wounds of conspecifics (Boesch 1992), which is so critical for healing that adult male macaques injured during attempts to enter a new group often temporarily return to their native group, where they are more likely to receive this service (Dittus & Ratnayeke 1989).

LEVELS OF EMPATHY

Emotional Contagion

The lowest common denominator of all empathic processes is that one party is affected by another's emotional or arousal state. This broad perspective on empathy, which goes back as far as Lipps (1903), leads one to recognize continuity between humans and other animals as well as between human adults and young children. Emotional connectedness in humans is so common, starts so early in life (e.g., Hoffman 1975, Zahn-Waxler &

Radke-Yarrow 1990), and shows neural and physiological correlates (e.g., Adolphs et al. 1994, Decety & Chaminade 2003a, Rimm-Kaufman & Kagan 1996) as well as a genetic substrate (Plomin et al. 1993), that it would be strange indeed if no continuity with other species existed. Evolutionary continuity between humans and apes is reflected in the similarity of emotional communication (Parr & Waller 2007) as well as similar changes in brain and peripheral skin temperature in response to emotionally charged images (Parr 2001, Parr & Hopkins 2001).

A flock of birds taking off all at once because one among them is startled shows a reflex-like, highly adaptive spreading of fear that may not involve any understanding of what triggered the initial reaction. Similarly, when a room full of human newborns bursts out crying because one among them started to cry, there is an automatic spreading of distress (Hoffman 1975). At the core of these processes is adoption—in whole or in part—of another’s emotional state, i.e., emotional contagion (Hatfield et al. 1993). Emotional contagion is not always a passive process, though: The object often aims to emotionally affect the subject, such as the extremely noisy temper tantrums of young apes when they are being rejected during weaning. Like human children (Potegal 2000), they exploit emotional contagion to induce maternal distress, which in turn may lead the mother to change her behavior to their advantage.

Emotional responses to displays of emotion in others are so commonplace in animals (de Waal 2003, Plutchik 1987, Preston & de Waal 2002b) that Darwin (1982 [1871, p. 77]) already noted that “many animals certainly sympathize with each other’s distress or danger.” For example, rats and pigeons display distress in response to perceived distress in a conspecific, and temporarily inhibit conditioned behavior if it causes pain responses in others (Church 1959, Watanabe & Ono 1986). A recent experiment demonstrated that mice perceiving other mice in pain intensify

their own response to pain (Langford et al. 2006).

Miller et al. (1959) published the first of a series of pioneering studies on the transmission of affect in rhesus macaques. These monkeys tend to terminate projected pictures of conspecifics in a fearful pose even more rapidly than negatively conditioned stimuli. Perhaps the most compelling evidence for emotional contagion came from Wechkin et al. (1964) and Masserman et al. (1964), who found that monkeys refuse to pull a chain that delivers food to them if doing so delivers an electric shock to and triggers pain reactions in a companion. Whether their sacrifice reflects concern for the other (see below) remains unclear, however, as it might also be explained as avoidance of aversive vicarious arousal.

Sympathetic Concern

The next evolutionary step occurs when emotional contagion is combined with appraisal of the other’s situation and attempts to understand the cause of the other’s emotions. De Waal (1996) speaks of “cognitive empathy” when the empathic reaction includes such contextual appraisal.

The psychological literature distinguishes sympathy from personal distress, which in their social consequences are each other’s opposites. Sympathy is defined as “an affective response that consists of feelings of sorrow or concern for a distressed or needy other (rather than sharing the emotion of the other). Sympathy is believed to involve an other-oriented, altruistic motivation” (Eisenberg 2000, p. 677). Personal distress, on the other hand, makes the affected party selfishly seek to alleviate its own distress, which mimics that of the object. Personal distress is not concerned, therefore, with the other (Batson 1991). A striking nonhuman primate example is how the continued screams of a punished infant rhesus monkey will cause other infants to embrace, mount, or even pile on top of the victim. Thus, one infant’s distress spreads quickly to its peers, which then seek to reduce

Sympathetic

concern: concern about another’s state and attempts to ameliorate this state (e.g., consolation)

Cognitive empathy:

empathy combined with contextual appraisal and an understanding of what caused the object’s emotional state

Personal distress:

self-centered distress born from empathy with another’s distress

Consolation:

comforting behavior directed at a distressed party, such as a recent victim of aggression

their own negative arousal (de Waal 1996, p. 46).

Concern for others is different in that it relies on a separation between internally and externally generated emotions. This separation is observable in many mammals. In a study that sought to document children's responses to family members instructed to feign sadness (sobbing), pain (crying), or distress (choking), striking similarities emerged between the reactions of one-year-old children and pets, such as dogs and cats. The latter, too, showed comforting attempts, such as putting their head in the lap of the "distressed" person (Zahn-Waxler et al. 1984).

Yerkes (1925, p. 246) reported how his bonobo, Prince Chim, showed such concern for his sickly chimpanzee companion, Panzee, that the scientific establishment might reject

his claims: "If I were to tell of his altruistic and obviously sympathetic behavior towards Panzee I should be suspected of idealizing an ape." Ladygina-Kohts (2001 [1935]) noticed similar tendencies in her young home-reared chimpanzee. She discovered that the only way to get him off the roof of her house (better than reward or threat of punishment) was by acting distressed, hence by inducing concern for herself in him.

Perhaps the best-documented example of sympathetic concern is consolation, defined as reassurance provided by an uninvolved bystander to one of the combatants in a previous aggressive incident (de Waal & van Roosmalen 1979). For example, a third party goes over to the loser of a fight and gently puts an arm around his or her shoulders (**Figure 1**). De Waal & van Roosmalen (1979) analyzed



Figure 1

Consolation is common in humans and apes, but virtually absent in monkeys. Here a juvenile chimpanzee puts an arm around a screaming adult male, who has just been defeated in a fight. Photograph by the author.

hundreds of consolations in chimpanzees, and de Waal & Aureli (1996) included an even larger sample. These studies show that bystanders contact victims of aggression more often than they contact aggressors, and bystanders contact victims of serious aggression more often than they contact those who had received mild aggression.

Subsequent studies have confirmed consolation in captive apes (Cordoni et al. 2004; Fuentes et al. 2002; Koski & Sterck 2006; Mallavarapu et al. 2006; Palagi et al. 2004, 2006), wild chimpanzees (Kutsukake & Castles 2004, Wittig & Boesch 2003), large-brained birds (Seed et al. 2007), and human children (Fujisawa et al. 2006). However, when de Waal & Aureli (1996) set out to apply the same observation protocol to detect consolation in monkeys, they failed to find any, as did others (Watts et al. 2000). The consolation gap between monkeys and the Hominoidea (i.e., humans and apes) extends even to the one situation where one would most expect consolation to occur: Macaque mothers fail to comfort their own offspring after a fight (Schino et al. 2004). O'Connell's (1995) content analysis of hundreds of reports confirms that reassurance of distressed others is typical of apes yet rare in monkeys. It still needs to be established, however, that this behavior actually does reduce the distressed party's arousal.

Empathic Perspective-Taking

Psychologists usually speak of empathy only when it involves perspective-taking. They emphasize understanding of the other, and adoption of the other's point of view. In this view, then, empathy is a cognitive affair dependent on imagination and mental state attribution, which may explain the skepticism about nonhuman empathy (Hauser 2000, Povinelli 1998). Perspective-taking by itself is, of course, hardly empathy: It is so only in combination with emotional engagement. The latter here is called "empathic perspective-taking," such as in one of the oldest

and best-known definitions by Smith (1759, p. 10) "changing places in fancy with the sufferer."

Menzel (1974) was the first to investigate whether chimpanzees understand what others know, setting the stage for studies of nonhuman theory-of-mind and perspective-taking. After several ups and downs in the evidence, current consensus seems to be that apes, but probably not monkeys, show some level of perspective-taking both in their spontaneous social behavior (de Waal 1996, 1998 [1982]) and under experimental conditions (Bräuer et al. 2005; Hare et al. 2001, 2006; Hirata 2006; Shillito et al. 2005).

A major manifestation of empathic perspective-taking is so-called targeted helping, which is help fine-tuned to another's specific situation and goals (de Waal 1996). The literature on primate behavior leaves little doubt about the existence of targeted helping, particularly in apes (see From Empathy to Altruism, below). A mother ape who returns to a whimpering youngster to help it from one tree to the next—by swaying her own tree toward the one the youngster is trapped in and then drape her body between both trees—goes beyond mere concern for the other. Her response likely involves emotional contagion (i.e., mother apes often briefly whimper themselves when they hear their offspring do so), but adds assessment of the specific reason for the other's distress and the other's goals. Tree bridging is a daily occurrence in orangutans, with mothers regularly anticipating their offspring's needs (van Schaik 2004, p. 104).

For an individual to move beyond being sensitive to others toward an explicit other-orientation requires a shift in perspective. The emotional state induced in oneself by the other now needs to be attributed to the other instead of the self. A heightened self-identity allows a subject to relate to the object's emotional state without losing sight of the actual source of this state (Hoffman 1982, Lewis 2002). The required self-representation is hard to establish independently, but one common avenue is to gauge reactions to a mirror.

Empathic perspective-taking: the capacity to take another's perspective—e.g., understanding another's specific situation and needs separate from one's own—combined with vicarious emotional arousal

Targeted helping: help and care based on a cognitive appreciation of the other's specific need or situation
