Altruism – A Philosophical Analysis

Christine Clavien, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

Michel Chapuisat, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

Introductory article

Article Contents

- Introduction
- Altruism in Biology
- Reproductive versus Psychological Altruism
- The Relationship between Reproductive and Psychological Altruism
- Other Forms of Human Altruism
- Conclusion

Online posting date: 15th March 2012

Altruism is a malleable notion that is understood differently in various disciplines. The common denominator of most definitions of altruism is the idea of unidirectional helping behaviour. However, a closer examination reveals that the term altruism sometimes refers to the outcomes of a helping behaviour for the agent and its neighbours – that is, reproductive altruism – and sometimes to what motivates the agent to help others – that is, psychological altruism. Since these perspectives on altruism are crucially different, it is important to use a clear terminology to avoid confusion. In particular, we show that the notion of altruism used by biologists profoundly differs from the ones used by philosophers, psychologists and economists in cross-disciplinary debates about human altruism.

Introduction

When a single term covers different concepts in diverse disciplines and when scholars from these disciplines read each other's papers, major confusions are to be expected. Mother Teresa's self-abnegation for the sake of the poor of Calcutta, and the reproductive restraint of ant workers are obviously distinct phenomena, yet both are described as altruistic (see Figure 1, Figure 2 and Figure 3). Mother Teresa's altruism refers to the common-sense use of the term – she is motivated by other-directed goals – whereas ants' altruism refers to a technical term well-defined in biology and independent of any moral consideration. Hence, ants qualify for altruism, as does Mother Teresa, but the two forms of altruism are not equivalent.

This chapter provides the conceptual tools necessary for understanding and evaluating the original contribution of various research areas involved in the project of explaining

eLS subject area: Bioethics & Philosophy

How to cite:

Clavien, Christine; and Chapuisat, Michel (March 2012) Altruism – A Philosophical Analysis. In: eLS. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0003442.pub2

the broad phenomenon of unidirectional helping behaviour. The section 'Altruism in Biology' deals with the way biologists define and explain altruism. The section 'Reproductive versus Psychological Altruism' is an introduction to altruism as conceived in philosophy and the social sciences. The section 'The Relationship between Reproductive and Psychological Altruism' deals with the relationship one can build between these two main forms of altruism. The section 'Other Forms of Human Altruism' draws attention to the ambiguous use of the term altruism in recent anthropology and economics literature on human altruism.

Altruism in Biology

In biology, altruism is generally defined as a behaviour that increases other organisms' fitness and permanently decreases the actor's own fitness (Frank, 1998; Grafen, 1985; West *et al.*, 2007). We label this form of altruism 'reproductive altruism'. Fitness is defined as the number of an actor's offspring that survive to adulthood (Hamilton, 1964: 1). In this acceptation, fitness is an absolute value – for example, a meerkat with four offspring at the end of her life has a fitness of four. This value however, only makes sense when compared to the fitness of other individuals in the population – for example, four offspring might be a great, medium or low reproductive achievement, depending on how many offspring meerkats usually have.

Classical examples of 'reproductive altruism' are found in the worker castes of eusocial insects such as termites, ants, some bees and some wasps. In these species, the workers are mostly sterile – their direct fitness is zero – and they raise the offspring of the queens and males. Altruistic castes have been reported in several other invertebrates, including gall-making aphids and thrips, snapping shrimps and flatworms, as well as in two species of mole rats. Such permanently sterile castes should be seen as the endpoint of a continuum. Less pronounced forms of reproductive altruism and cooperation are widely distributed, from bacteria and amoebae to cooperatively breeding birds and mammals. See also: Eusociality and Cooperation



Figure 1 Altruism in ants: workers of *Formica selysi* groom the queen. Photo courtesy of David Buchs, http://www.lesixiemecontinent.net/



Figure 2 Altruism in ants: workers of *Formica selysi* take care of the queen's offspring. Photo courtesy of Joël Meunier.

Reproductive altruism is an extreme form of helping behaviour, the evolution of which cannot be accounted for in terms of future reciprocity (Trivers, 1971) or individual reputation (Zahavi, 1975). It poses a challenge for evolutionary theorists because natural selection should favour traits increasing the reproduction of their bearers. Since altruistic behaviours are, by definition, detrimental with respect to survival and reproduction, their persistence



Figure 3 A memorial plate for Mother Teresa, an icon of human altruism. Photo by Michal Manas. http://commons.wikimedia.org/wiki/File:Mother_Teresa_memorial_plaque.jpg

in the course of evolution calls for a special explanation. **See also**: Theory of Cooperation

After several unsatisfying explanatory attempts (e.g. Darwin, 2007/1871; Lorenz, 2004/1963; Wynne-Edwards, 1962), William Hamilton (1964, 1970, 1975) provided a powerful explanation for reproductive altruism: kin selection theory. Hamilton generalised and formalised an idea already alluded to by Fisher (1958) and Haldane (1955). The latter famously said that he would jump into a river for two brothers or eight cousins. The idea is that reproductive altruism can be favoured when the altruists help genetically related individuals, which carry identical copies of genes. In a gene-centred formulation, genes responsible for altruistic behaviour can spread in a population if the altruistic behaviour they induce is more likely to benefit organisms who possess copies of the same altruistic genes (Frank, 1998; Grafen, 1985; Hamilton, 1970; Queller, 1992; West et al., 2011). Such a situation typically occurs among close relatives, due to common ancestry. For example, altruists may preferentially help their parents, sisters or cousins, based on their ability to recognise close kin or simply because they live in spatial proximity.

It is worth noting however, that common ancestry is only one way to increase the probability that the altruist and her recipient share the gene – or complex of genes – for altruism. Another way for altruists and recipients to share the altruism gene is the so-called 'green beard' effect, whereby a gene for altruism is responsible for a visible phenotype – a green beard or anything more realistic – that can be recognised by other altruists. If in addition, the tendency for altruism is coupled with a preference for helping those who share the visible phenotype, common ancestry is not

needed for altruism to spread in the population. In the natural world, green beard cases have been reported (Keller and Ross, 1998) but they are rare because helping otherwise unrelated altruists is wasteful for all the genes of the organisms except for the altruistic gene; thus, any mutation that prevents this waste will be favoured by natural selection (Dawkins, 1979; Maynard Smith, 1976; Okasha, 2002). See also: Eusociality and Cooperation; Hamilton, William Donald; Theory of Cooperation

Hamilton formulated a simple rule delineating when reproductive altruism can evolve by kin selection: rb-c>0, where r is the coefficient of genetic relatedness between the altruistic actor and the recipient — which reflects the probability that both individuals carry the gene for altruism, compared to the population average (Frank, 1998), b the fitness benefit of the altruistic behaviour for the recipient and c the fitness cost of the altruistic behaviour for the actor. When the inequation is fulfilled, the altruist transmits more copies of his genes by helping his relative (rb) than by reproducing (c) and therefore reproductive altruism can evolve.

In the literature, one can find an alternative way to model the evolution of altruism. The 'group selection' approach (Goodnight, 1990; Odling-Smee et al., 2003; Sober and Wilson, 1998; Wade, 1977; Wilson, 1975) considers groups as functional units and explicitly partitions selection into a between-group and a within-group component. In some circumstances, it might happen that an altruistic behaviour is so favourable at the group level that it compensates for the loss of individual fitness within groups. Interestingly, even if the two methods look at the same phenomenon from different perspectives, they rely on similar quantitative genetic approaches and yield fully compatible results (Gardner, 2008; Lehmann et al., 2007). Using Price's (1970) equation, several authors have demonstrated mathematically that the change of frequency of an altruistic gene in a population can be equivalently expressed either in terms of Hamilton's rule or in terms of a partition of selection within and between groups. In particular, the two approaches concur in the view that true reproductive altruism, with a permanent cost to the actor, will only evolve in groups of genetically related individuals (e.g. Frank, 1998; Gardner et al., 2007; Lehmann and Keller, 2006; Queller, 1992). See also: Group Selection; Selection: Units and Levels

In practical cases, it might be difficult to distinguish true reproductive altruism from cooperation. Indeed, the long-term individual costs and benefits of a given behaviour are often difficult to measure. Many helping behaviours that might seem altruistic at first may provide delayed direct fitness benefit. For example, vampire bats have been observed to feed their offspring as well as other individuals of their group (Wilkinson, 1984). Under close observation, it turns out that the latter pay back when they return from a successful hunting night. This behaviour is more correctly described as cooperative food sharing. Many behaviours might also be altruistic in some circumstances and cooperative in others. Finally, apparently altruistic

behaviours may in fact confer delayed direct fitness benefits through reputation or prestige. Helping behaviours may indeed increase the individual reputation and be honest signals, for example, for being a good mate or a reliable partner for future cooperative interactions (Zahavi, 1975).

Reproductive altruism should not be confused with two other notions found in the literature: 'reciprocal altruism' and 'weak altruism'. In an important and influential paper, Trivers (1971) discussed reciprocal behaviour of the form 'help your neighbour as long as she helps you back'. He labelled this behaviour 'reciprocal altruism', while making it clear that it provides a long-term direct benefit for its bearer, and as such does not exemplify altruism in the classical Hamiltonian sense defined above. Reciprocal altruism is a form of cooperative behaviour. One illustration is the vampire bats food sharing example mentioned above.

Proponents of group selection theory often use a slightly different notion of altruism, labelled 'weak altruism' (Wilson, 1977). Weak altruism is not an absolute but a relative value that compares the actor's number of offspring to that of the recipients. To qualify as weakly altruistic, a helping individual must be less fit than the group members who do not help, but benefit from help. Hence, this definition does not imply that altruistic individuals reduce their direct fitness. To illustrate this point, consider the case of an individual who builds a protective wall around a common nest. The wall is costly to build but this cost is compensated by the direct benefits it provides to the individual – say protection against floods. In this case, the wall-building behaviour is not 'reproductively altruistic' because the individual cost is compensated by the benefit, but is 'weakly altruistic' – because the other nest members benefit from the wall without having to build it. These conceptual distinctions allow us to solve some apparent contradictions among authors. For example, in two famous papers, Wilson (1975) and Hamilton (1975) built very similar models for the evolution of a trait that Wilson described as altruistic and Hamilton as nonaltruistic. In fact, Wilson had 'weak altruism' in mind whereas Hamilton referred to reproductive altruism. See also: Group Selection; Selection: Units and Levels

These conceptual distinctions are also important to avoid cross-disciplinary misunderstandings. In the humanities and social sciences, scholars are often unaware of the various forms of altruism used in biology. They frequently mistake reproductive altruism for reciprocal altruism or psychological altruism. After having heard of Trivers' theory, they infer that biologists deny the possibility of any form of altruism that would not be individually advantageous. This confusion is usually coupled with misunderstandings about the level of selection. Indeed, Hamilton's explanation of the conditions under which genes for altruism can be *successfully* passed on from one generation to the next has been widely popularised by Richard Dawkins (1976) in the terminology of the 'selfish gene'. This metaphorical use of the term 'selfish' for genes simply emphasises the fact that genes are the basic unit of selection. Dawkins made it clear that 'selfishness' at the genetic level did not imply selfishness at the social, individual or psychological levels. However, this distinction is widely overlooked by scholars who are not trained in population genetics. Consequently, they erroneously think (see for example Gintis *et al.*, 2005; Sahlins, 1976) that Hamilton and his followers, by adopting the gene-centred point of view that is now standard in evolutionary biology, deny the existence of any form of altruism.

Reproductive versus Psychological Altruism

Evolutionary theorists define altruism in terms of its consequences on the fitness of individuals. This focus on the outcomes of a behaviour, however, does not correspond to the common-sense use of the term altruism, which is more concerned with what *motivates* people to help. If we were to learn that Mother Teresa became a nun and served the poor on the streets of Calcutta only because she feared going to Hell, we would hardly describe her as an altruist. Conversely, a father might ruin himself to finance the legal studies of his son with the genuine hope of providing him with bright career opportunities and a happy future life, but if his son's secret dream is to become a ballet dancer, the father's altruistic motive might fail to achieve the purported goal. Despite the father's erroneous judgement about what will make his son happy, common sense will appreciate the altruistic nature of his motivation, as long as it is not tainted by self-directed calculations.

An important and longstanding cross-disciplinary debate is based on this common-sense understanding of altruism. The debate is about deciding whether humans are capable of 'psychological altruism', which, as the psychologist Daniel Batson describes it, is 'a motivational state with the ultimate goal of increasing another's welfare'. Under this acceptation, altruism refers to an 'energy, a force within the individual' which drives the person to achieve an other-directed goal (Batson, 1991: 6). Psychological altruism is mainly discussed in philosophical circles (Butler, 1991/1726; Hutcheson, 2004/1725; Mandeville, 1997/1714–1728; Rand, 1964; Stich *et al.*, 2010) as well as in the social sciences more generally, for example, psychology (Batson, 1991; Cialdini *et al.*, 1987) and sociology (Comte, 1851–1854).

In order to avoid confusion, it is important to understand that reproductive altruism and psychological altruism are two distinct notions (Sober and Wilson, 1998). Ant workers are paradigmatic examples of the former, but their limited cognitive abilities do not permit them to conceptualise others' welfare. Thus, it does not even make sense to ask the question whether they are psychological altruists. Conversely, mothers caring for their offspring are not reproductive altruists although they might be genuinely motivated to promote their children's well-being.

The Relationship between Reproductive and Psychological Altruism

The distinction between ultimate and proximate explanations is central to evolutionary biology (Mayr, 1961). *Ultimate* explanations refer to the adaptive value and fitness consequences of a trait; they answer the question of why a trait has evolved. *Proximate* explanations provide details about the causal mechanisms responsible for a trait. These explanations are distinct but complementary. Biologists rely on both to fully comprehend a phenomenon. For example, kin selection elucidates why soldier ants defend their colony against intruders threatening their mother and relatives – an ultimate explanation – whereas other studies provide details on how these soldiers detect intruders thanks to their sensitivity to chemical cues – a proximate explanation.

The same distinction applies to the two ways of conceiving altruism presented in the previous sections. The Hamiltonian or group selection explanations for the evolution of reproductive altruism are *ultimate* explanations for this behaviour. In contrast, psychological altruism refers to motivations that are causal mechanisms prompting individuals to take care of others' welfare; this psychological phenomenon provides a *proximate* explanation for the broad phenomenon of unidirectional helping behaviour.

It should be clear, however, that the demonstration of an individual's altruistic motivation – proximate explanation - does not indicate that this individual is reproductively altruistic. For example, the individual may care for his own children. Conversely, an ultimate evolutionary explanation of reproductive altruism provides no direct insight into what motivates individuals to help others – for example, we do not know what worker ants think. There are however, interesting indirect links between the two types of altruism. If a form of reproductive altruism has been selected in humans, there are good reasons to think that some proximate mechanisms prompting this fitness-costly behaviour have co-evolved. Such proximate mechanisms might be caring and empathic emotions, which may qualify as psychological altruism (Clavien, 2012; Sober and Wilson, 1998). More generally, mechanisms based on psychological altruism such as caring and empathic emotions are effective ways to make individuals behave altruistically or cooperatively, depending on whether or not these behaviours have long-term direct fitness costs to the individual. Thus, in highly cooperative species in which individuals are endowed with developed cognitive and emotional capacities – such as humans and probably apes – various forms of psychological altruism are likely to be selected (Clavien, 2010: Kitcher, 2011).

Other Forms of Human Altruism

More recently, the notion of altruism has been heavily used in the emerging fields of experimental economics and evolutionary anthropology (de Quervain et al., 2004; Fehr and Fischbacher, 2003; Gintis et al., 2005; Henrich et al., 2005: 812). This literature provides empirical data – gathered in laboratory experiments - that highlight humans' cooperative and fair behaviour. In these studies, the meaning of the term altruism is variable and often ambiguous (West et al., 2011). Sometimes, altruism refers to a particular type of behaviour that can be observed in socio-economic contexts, much in the same spirit as the biological use of the term. At other times, altruism refers to people's willingness to cooperate and punish shirkers, which is close to the common-sense use of the term. However, a closer look at the literature shows that the authors often use more fuzzy and less restrictive notions of altruism than the ones presented above. Their outcomeoriented notion is less restrictive than reproductive altruism because they count any type of individual success – rather than exclusively reproductive success – and show little interest in the broad evolutionary impact of the behaviours they observe in laboratory. Their motivationoriented notion refers to preferences, as defined in neoclassical economic theory, which are weak indicators of people's actual motivations. Overall, the results obtained in these experiments are important for understanding cooperation, but not reproductive altruism.

Conclusion

We have specified two main understandings of altruism. The first concentrates on the fitness outcomes of a behaviour, whereas the second focuses on the individual motivation underlying an action. These two notions of altruism are distinct but complementary ways of shedding light on unidirectional helping behaviour. We have also hinted at the fact that other concepts of altruism pervade current literature on human altruism. We conclude that a true interdisciplinary understanding of human altruism crucially relies on distinguishing the multifarious notions covered by the term altruism in diverse disciplines.

References

- Batson CD (1991) The Altruism Question: Toward a Social Psychological Answer. Hillsdale: Lawrence Erlbaum.
- Butler J (1991) Fifteen sermons. In: Raphael DD (ed.) *British moralists, 1650–1800: Selected and Edited with Comparative Notes and Analytical Index*, vol. 1, pp. 325–377. Oxford: Clarendon Press.
- Cialdini RB, Schaller M, Houlihan D et al. (1987) Empathy-based helping: is it selflessly or selfishly motivated? Journal of Personality and Social Psychology 52(4): 749–758.
- Clavien C (2010) Je t'aide moi non plus: biologique, comportemental ou Psychologique, l'altruisme dans tous ses états. Paris: Vuibert.
- Clavien C (2012) Altruistic emotional motivation: an argument in favour of psychological altruism. In: Plaisance K and Reydon T (eds) *Philosophy of Behavioral Biology*. Boston Studies in Philosophy of Science, pp. 275–296. Dordrecht: Springer Press.

- Comte A (1851–1854) Système de politique positive, ou, Traité de sociologie instituant la religion de l'humanité. Paris: L. Mathias.
- Darwin C (2007) The Descent of Man, and Selection in Relation to Sex (Trans: Zimmer C), Concise edn. New York: Plume.
- Dawkins R (1976) *The Selfish Gene*. New York: Oxford University Press.
- Dawkins R (1979) 12 Misunderstandings of Kin selection. *Zeits-chrift Fur Tierpsychologie Journal of Comparative Ethology* **51**(2): 184–200.
- Fehr E and Fischbacher U (2003) The nature of human altruism. *Nature* **425**(6960): 785–791.
- Fisher RA (1958) *The Genetical Theory of Natural Selection*. Dover Books on Science, vol. S466, 2nd rev. edn. New York: Dover Publications.
- Frank SA (1998) Foundations of Social Evolution. Monographs in Behavior and Ecology. Princeton, NJ: Princeton University Press.
- Gardner A (2008) The price equation. *Current Biology* **18**(5): R198–R202.
- Gardner A, West SA and Barton NH (2007) The relation between multilocus population genetics and social evolution theory. *American Naturalist* **169**(2): 207–226.
- Gintis H, Bowles S, Boyd R and Fehr E (eds) (2005) Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life. Cambridge, MA: MIT Press.
- Goodnight CJ (1990) Experimental studies of community evolution (II & II). *Evolution* **44**(6): 1614–1636.
- Grafen A (1985) A geometric view of relatedness. In: Dawkins R and Ridley M (eds) *Oxford Surveys in Evolutionary Biology*, vol. 2, pp. 905–907. Oxford: Oxford University Press.
- Haldane JBS (1955) Population genetics. *New Biology* **18**: 34–51. Hamilton WD (1964) The genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**(1): 1–52.
- Hamilton WD (1970) Selfish and spiteful behaviour in an evolutionary model. *Nature* **288**(5277): 1218–1220.
- Hamilton WD (1975) Innate social aptitudes of man; an approach from evolutionary genetics. In: Fox R (ed.) *Biosocial Anthropology*, pp. 133–155. London: Malaby Press.
- Henrich J, Boyd R, Bowles S *et al.* (2005) 'Economic man' in crosscultural perspective: behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences* **28**(6): 795–855.
- Hutcheson F (2004) An Inquiry into the Original of Our Ideas of Beauty and Virtue. Indianapolis, IN, USA: Liberty Fund.
- Keller L and Ross KG (1998) Selfish genes: a green beard in the red fire ant. *Nature* **394**(6693): 573–575.
- Kitcher P (2011) *The Ethical Project*. Cambridge, MA: Harvard University Press.
- Lehmann L and Keller L (2006) The evolution of cooperation and altruism: a general framework and a classification of models. *Journal of Evolutionary Biology* **19**(5): 1365–1376.
- Lehmann L, Keller L, West SA and Roze D (2007) Group selection and kin selection: two concepts but one process. *Proceedings of the National Academy of Sciences of the USA* **104**(16): 6736–6739.
- Lorenz K (ed.) (2004) Das sogenannte Böse: zur Naturgeschichte der Aggression, dtv 33017, 24. Aufl. edn. München: Deutscher Taschenbuch Verl.
- Mandeville B (1997) *The Fable of the Bees: And Other Writings*. Indianapolis: Hackett Publishing compagny.
- Maynard Smith J (1976) Group selection. *The Quarterly Review of Biology* **51**(2): 277–283.

- Mayr E (1961) Cause and effect in biology. *Science* **134**: 1501–1506.
- Odling-Smee FJ, Laland KN and Feldman MW (2003) *Niche Construction: The Neglected Process in Evolution.* Monographs in Population Biology, vol. 37. Princeton, NJ: Princeton University Press.
- Okasha S (2002) Genetic relatedness and the evolution of altruism. *Philosophy of Science* **69**(1): 138–149.
- Price GR (1970) Selection and covariance. *Nature* **227**(5257): 520–521.
- Queller DC (1992) A general-model for kin selection. *Evolution* **46**(2): 376–380.
- de Quervain DJF, Fischbacher U, Treyer V *et al.* (2004) The neural basis of altruistic punishment. *Science* **305**(5688): 1254–1258
- Rand A (1964) The Virtue of Selfishness: A New Concept of Egoism. New York: New American Library.
- Sahlins MD (1976) *The Use and Abuse of Biology: An Anthropological Critique of Sociobiology.* Ann Arbor, MI: University of Michigan Press.
- Sober E and Wilson DS (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stich SP, Doris JM and Roedder E (2010) Altruism. In: Doris JM (ed.) *The Moral Psychology Handbook*, pp. 147–205. Oxford: Oxford University Press.
- Trivers RL (1971) The evolution of reciprocal altruism. *The Quarterly Review of Biology* **46**(1): 35–57.

- Wade MJ (1977) Experimental study of group selection. *Evolution* **31**(1): 134–153.
- West SA, El Mouden C and Gardner A (2011) Sixteen common misconceptions about the evolution of cooperation in humans. Evolution and Human Behavior 32(4): 231–262.
- West SA, Griffin A and Gardner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* **20**: 415–432.
- Wilkinson GS (1984) Reciprocal food sharing in the vampire bat. *Nature* **308**(5955): 181–184.
- Wilson DS (1975) A theory of group selection. *Proceedings of the National Academy of Sciences of the USA* **72**(1): 143–146.
- Wilson DS (1977) Structured demes and the evolution of group-advantageous traits. *American Naturalist* **111**(977): 157–185.
- Wynne-Edwards VC (1962) Animal Dispersion in Relation to Social Behaviour. New York: Hafner.
- Zahavi A (1975) Mate selection a selection for a handicap. *Journal of Theoretical Biology* **53**(1): 205–214.

Further Reading

- Foster KR (2008) Altruism. In: Jorgensen SE and Fath B (eds) *Encyclopedia of Ecology*, pp. 154–159. Amsterdam: Elsevier.
- Keller L and Chapuisat M (1999) Cooperation among selfish individuals in insect societies. *BioScience* **49**(11): 899–909.
- Lehmann L, Bargum K and Reuter M (2006) An evolutionary analysis of the relationship between spite and altruism. *Journal of Evolutionary Biology* **19**(5): 1507–1516.