

How Folk Psychology has Determined Evolutionary Theories About Altruism: An Alternative Perspective Based on Buddhist Theory

Robert L. West (robert_west@carleton.ca)

Institute of Cognitive Science, Department of Psychology,
Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada

Abstract

This paper explores how commonly held folk psychology beliefs about selfishness and altruism have framed theoretical accounts of whether, or how, altruism could have evolved. This is demonstrated by assuming an alternative view of selfishness and altruism (from the Tibetan/Buddhist culture) and showing that it results in a very different, but completely plausible account of altruism and selfishness within evolutionary theory.

Keywords: Evolution; altruism; selfishness; folk psychology; Buddhism.

The question about what the theory of Evolution implies about the human ability to behave altruistically is a controversial one. Popular, mainstream evolutionary theorists (e.g., Dennett, 2007; Dawkins, 2007) argue that the theory of Evolution implies we are fundamentally selfish, with acts of apparent altruism being explained through the mechanisms of reciprocity (Trivers, 1971) and kin selection (Haldane, 1955; Hamilton, 1963). The only exception to this is the recent resurfacing of the theory of altruism through group selection (Sober & Wilson, 1998; Wilson, 2003), which argues that genuine altruism could have evolved by providing an advantage at the level of group competition. However, what all of these arguments have in common is that they implicitly frame the discussion of this issue using a particular model of how selfishness and altruism operate in the human brain. This model is not scientifically derived; instead it is based the folk psychology of our modern, western culture. Furthermore, this folk psychology model of selfishness and altruism is not culturally universal. In this paper I argue that this debate, so far, has been unknowingly constrained and determined by the assumed folk psychology model. I demonstrate this by recasting the issue in terms of the Tibetan/Buddhist model of selfishness and altruism, and showing that it produces a very different but equally plausible view of altruism within evolutionary theory.

At first glance it seems that organisms must be constructed around an inherent sense of selfishness, since the alternative (constructing them around altruism) doesn't make evolutionary sense. After all, if I don't survive to reproduce then my genes are gone. As noted above, the only exception to this is based on the idea of group selection. However, groups need to cooperate in some way to enhance fitness before group selection can occur. Therefore, group selection can occur only after organisms have evolved sufficiently to begin to produce some form of beneficial

cooperative behavior. Thus the standard way of framing this issue is to cast the earliest creatures as selfish and place the burden of evolving on altruism (i.e., through the process of group selection). This is reflected in the fact that nobody ever asks how selfishness could have evolved. However, this way of framing things is based on the assumption that the dichotomous division of behavior into or selfish or altruistic is valid. If it is valid then two things follow: (1) the first creatures would have been selfish and (2) altruism had to come into being against the backdrop of a completely selfish world. All current theories about evolution and altruism make these two assumptions.

Dividing behavior into selfish (benefiting yourself) and altruistic (benefiting another at a cost to yourself) is valid when discussing the process of evolution because these terms have a very precise meaning within the theory of evolution. The problem arises when this dichotomy is extended to characterize the internal motivational states of the creatures that are evolving. That is, when it is somehow assumed that the selfishness required for surviving and reproducing is actually in the creature. This is because evolutionary success does not flow from the creature, but rather from the interaction between the environment and the creature. The idea that the outcome of the interaction between the agent and the environment is not always a direct reflection of the motivations or cognitions of the agent has been explored and developed through a number of theoretical positions – e.g., Gibsonian Affordances (e.g., Gibson, 1979), Situated Cognition (e.g., Robbins & Aydede, 1994), Distributed Cognition (e.g., Hutchins, 1994), and Dynamic Systems Theory (e.g., Port, & vanGelder, 1998). This general principle is commonly accepted for explaining animal behavior. For example, ants do not intend to build ant colonies and birds do not plan to flock. Instead these outcomes are emergent properties that arise from other motivations. Likewise, it is possible for selfish behaviors (in the evolutionary sense) to arise without selfish intentions. The group selection argument is based on this general idea (i.e., people are motivated to sacrifice themselves for the good of the group, but this unwittingly increases the likelihood of passing on their genes).

In addition to being familiar to us from its use within the theory of Evolution, the altruism/selfishness dichotomy also seems familiar because having only two states is consistent with the modern western folk psychology view of selfishness and altruism. Here, by folk psychology, I mean a theory about how the mind works, shared among a group, that is based on introspection and experience rather than a scientific or experimental approach. Also, I do not mean it in a pejorative

sense. A folk psychology theory could be inaccurate and misleading or it could be very close to the truth. The main point is that it is not based on scientific evidence.

Generally speaking, folk psychology theories can be understood in terms of the states of the mind that are postulated and the way in which these states interact with each other. The western folk psychology view is that there are two states, selfishness and compassion, and these interact by producing conflicting motivations (I will use compassion to refer to the motivational state that leads to altruism). In this folk psychology system, one state is not necessarily considered to be more fundamental than the other; beliefs about this depend on the philosophical and cultural views of the individual. Mainstream Evolutionary theory accepts the western folk psychology view in that it implicitly accepts that this issue should be framed using the two western folk psychology states to define the range of possible states. As noted above, this framework, in combination with the general principles of evolution, produces the conclusion that: (1) the first creatures would have been selfish and (2) altruism had to come into being against the backdrop of a completely selfish world.

However, this starting point is problematic because it is not culturally universal. One important exception lies in Tibetan Buddhist culture. In the Tibetan language they have three terms, rather than two, to describe helping behaviors (Goleman, 2004). Two of these terms are similar to what we know as altruism (i.e., the urge to help another) and selfishness (i.e., the urge to help yourself). To avoid confusion, when referring to selfishness and altruism in the Tibetan sense I will append a (T) for Tibetan and when referring to selfishness and altruism in the Western sense I will append a (W) for Western. The third term in the Tibetan system refers to a generalized urge to reduce suffering, which could lead you to help another or to help yourself, depending on where you perceive the suffering. The third term, which I will refer to as GC (for Generalized Compassion), does not result in a third type of outcome, you still either help yourself or help someone else, but the point is that it does not *feel* the same. That is, it is a label for a different motivational state (Goleman, 2004).

In addition to having three different terms in the Tibetan language. Tibetan Buddhism, and Buddhism in general, provides a relatively clear model of how these states interact. According to this view, the reason GC feels different is that selfishness(T) and altruism(T) are based on attachment to the object being helped, and GC is not. Attachment, in the Buddhist sense, means that you want things to be a certain way. Thus, altruism(T) can be viewed as consistent with mainstream evolutionary beliefs that the real motivation behind altruistic acts is to benefit the helper. For example, a mother who makes sacrifices so her child can go to medical school could be acting due to her attachment to the idea of having a doctor in the family. In contrast, GC refers to a general urge to reduce suffering, regardless of whether the suffering is in yourself or in another. For example, a mother who soothes her child when it cries could be acting based on

GC. However, she could also be acting based on altruism(T), with a desire to have a quiet child. Therefore, according to the Buddhist view, there are two systems acting in parallel, the attachment based altruism(T)/selfish(T) system and the GC system. However, it is important to note that, because these systems act in parallel, the resulting motivational state is normally considered to be a mixture of these sources (Goleman, 2004).

In the Tibetan case, although they have a word for GC in their language, it is unclear to me at this time how widespread the use of the GC concept is amongst ordinary Tibetans. However, this model is common to Buddhists. Evidence of similar intuitions can also be found in other groups with a similar focus on compassion. For example, the Greek term *agape*, which has largely fallen out of use, was previously used by Christians in the West to describe a motivational state very similar to GC. Also, there are references to a GC-like mechanism in various Christian monastic writings (e.g., see the 12th article of Fenelon's Maxims of the Saints), although these tend to be more convoluted because of the need to interpret the feeling of GC within a theistic framework. I will refer to the three state model as the Buddhist model because the Buddhist version of the model is very clear and not inconsistent with a scientific view of the mind, however, it is important to keep in mind that intuitions about the existence of GC are not limited to a particular group. For brevity I will refer to the modern western folk psychology model discussed above as the Western model, because it is currently common in the western industrialized countries.

In terms of understanding the motivation for helping behaviors, the Buddhist model and the Western model are both cases of folk psychology. That is, they are both based on introspection and experience. Neither has any special scientific status. The Western model is simpler and therefore some might argue it is more parsimonious. However, the parsimony standard should not be raised until the implications of the Buddhist system have been worked out (see below). In contrast, it can be argued that the Buddhist model is the product of greater efforts to introspect on compassion, and therefore could be more accurate. However, since we cannot, at this point, scientifically determine which model is correct, the main issue is whether or not the Buddhist model makes sense within an Evolutionary framework. That is, does it result in a plausible, parsimonious account of selfishness and altruism within the evolutionary process?

In terms of applying the Buddhist model to Evolution, the first thing to note is that the first creatures cannot be exclusively selfish, otherwise GC will run into the same evolutionary road blocks as altruism(W). Also, it is hard to think of a reason why GC would have evolved in this scenario. Therefore, it is necessary to postulate GC as the original motivational mechanism for the first creatures. This position will seem incredible to some readers (I know this from feedback), however, as outlined below, for very simple creatures there is no difference between acting according to selfishness(W) and acting according to GC.

For creatures with a limited ability to sense and act in the world, GC is completely sufficient to account for their behavior within an Evolutionary framework. This is because the only suffering they would be aware of is their own, so they would act only in their own best interests. Note also that operating according to GC under these conditions is simple and parsimonious – *if there is suffering (e.g., a need, a deviation from a healthy equilibrium, an external threat), then act to reduce it*. Thus, the story can begin with the first creatures operating only according to GC, which is at least as parsimonious as the first creatures operating only according to selfishness(W). In fact, the use of GC could be argued to be more parsimonious as it does not need to be scaled up to explain the selfishness of creatures that are aware of the suffering of others. GC is simply the urge to reduce suffering, which could be said to exist in any biological creature that needs to maintain an equilibrium to survive.

With GC as the starting point, the question is then, how did selfishness(T) and altruism(T) evolve? Operating according to GC only becomes an issue once a creature has evolved sufficiently to sense suffering in another creature. At this point GC could result in a creature spending resources to reduce the suffering of another creature. From an evolutionary point of view, this is a bad strategy if you help a competitor, but it could also be a good strategy if the other creature is your offspring or carries some of your genes or is willing to mate with you once their survival is assured.

In theory, different scenarios are possible for the development of selfishness(T) and altruism(T), but I will follow one that parallels the way selfishness(W) and altruism(W) are commonly understood within mainstream Evolutionary theory. As noted above, once a creature can sense the suffering of another, GC can provide both advantages and disadvantages. Therefore, selfishness(T) could have evolved as an opposing force to GC in the areas where GC produced an evolutionary disadvantage. Likewise altruism(T) could have evolved, through the mechanisms of reciprocity and kin selection, to favor helping behaviors that result in evolutionary advantages. However, it seems unlikely that selfishness(T) and altruism(T) would have replaced GC since all that was needed was to bias the system toward evolutionarily advantageous behaviors. Thus in this system, most actions would be a mixture, with GC providing a general urge to reduce suffering, and selfishness(T) and altruism(T) providing an increased weighting for actions that produce evolutionarily advantages. Note also that selfishness(T) and altruism(T) are not assumed at the beginning. Instead they are evolved according to straight-forward evolutionary pressures. Thus a plausible evolutionary story requires assuming only one motivational state (GC) at the beginning, the same as the dominant selfishness(W) view.

As we can see, by assuming the Buddhist model an alternative story emerges that offers a plausible account of altruism within evolution that is significantly different from

current, mainstream views. One of the most important differences is that there is no need to explain how altruism could have evolved in a completely selfish world. Instead the motivational mechanism for altruism (GC) predates selfishness, which evolves as a response to a greater awareness of the suffering of others. Group selection is no longer necessary to evolve the neural mechanisms for altruism from scratch. However, group selection may have preserved or even promoted broad based altruistic behaviors based on GC, since these would benefit the group as a whole. Overall, compared to the selfish(W) model, the GC model makes group selection for altruistic behaviors more likely because there is a preexisting basis for altruistic behaviors to occur.

Up until now, the debate about altruism and Evolution has been constrained by the implicit assumption that the Western folk psychology model is valid. This assumption is questionable. In fact, there is very little direct evidence for the validity of this model. The GC model, as presented in this paper, is an alternative theory, based on equally legitimate assumptions, that produces a plausible and parsimonious account of the relationship between altruism and evolution. Although we cannot, at this time, ascertain which theory is right, what this shows is that assumptions (implicit or explicit) about how selfishness and altruism operate in the human mind play a key role in theorizing about the relationship between altruism and Evolution.

References

- Dawkins, D. *The God Delusion*. Black Swan (2007)
- Dennett, D. C. *Breaking the Spell: Religion as a Natural Phenomenon*. Penguin Books (2007)
- Goleman, D. *Destructive Emotions: A Scientific Dialogue with the Dalai Lama*. Bantam; Reprint edition (2004).
- Gibson, James J.: *The Ecological Approach to Visual Perception*. New Jersey, USA, Lawrence Erlbaum Associates (1979).
- Haldane, JBS. 1955. Population Genetics. *New Biology* 18:34-51
- Hamilton, WD. 1963. The evolution of altruistic behavior. *American Naturalist* 97:354-356
- Hutchins, E.,. *Cognition in the wild*. MIT Press ISBN 0-262-58146-9(1995)
- Philip Robbins and Murat Aydede (Eds.), *Cambridge Handbook of Situated Cognition*. New York: Cambridge University Press.(1994)
- Port, R. F., & vanGelder, T. (Eds.) *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press (1998)
- Sober, E., & Wilson, D. S. *Unto Others: Evolution and Psychology of Unselfish Behavior*. Harvard University Press (1998)
- Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*. 46: 35-57.
- Wilson, D. S. *Darwin's Cathedral: Evolution, Religion and the Nature of Society* University of Chicago Press (2003)