**"Altruism and Biology"**
**Entry for *International Encyclopedia of Ethics* (Wiley-Blackwell, 2013)**
**Richard Joyce**

TITLE: altruism and biology
WORD COUNT: 3299

In the vernacular, "altruism" is a psychological term; its application depends on the motives with which an action is performed. A's action is altruistic if and only if A's ultimate motive in performing the action is the furtherance of some other individual's (or individuals') welfare. A's action remains altruistic even if, as a matter of fact, the other individual receives no benefit (because of some unforeseen interference, say). And A's action remains altruistic even if, as a matter of fact, A (and A alone) benefits from the action, so long as that outcome was not A's ultimate motive. If A acts to benefit B but A's *ultimate* motive for doing so is that furthering B's welfare promises in turn to benefit A (via boosting A's reputation, say), then A's action is not altruistic but selfish. On the face of it, there is no reason to assume that this distinction is exhaustive. It seems possible that actions can be done for motives that have nothing to do with benefiting anyone, or can be done for mixed motives where benefits to both actor and other are important motivating factors but neither is ultimate.

This last claim has, however, been denied. Some have argued that all human actions have the same ultimate motive: benefiting the actor. This view is called "psychological egoism." (*See* EGOISM.) Psychological egoism is sometimes thought of as almost trivially true, on the grounds that all human actions are intentional—meaning that the actor *desires* to perform the action—thus all actions satisfy desires, thus all actions bring the actor some satisfaction. But this argument is faulty on at least two grounds. First, the satisfaction of a desire need not bring satisfaction. As Joel Feinberg says: "Sometimes when we get what we want we *also* get, as a kind of extra dividend, a warm, glowing feeling of contentment; but often, far too often, we get no dividend at all, or, even worse, the bitter taste of ashes" (2008: p. 523). Second, and more decisively, even if it were true that all actions bring some satisfaction, it still would not follow that the gaining of this satisfaction is the ultimate motive for every action.

It is tempting to assume that psychological egoism is generally motivated by a pessimistic or cynical view of humans. But even the opponent of psychological egoism can allow that humans are *nearly always* selfish; it is enough to defeat egoism if altruistic actions occasionally occur. Moreover, a really pessimistic (though plausible) view of humans must allow for truly *spiteful* actions—actions done for the ultimate motive of harming another, where the actor acknowledges that she will be worse off as a result of performing the action. (We must bear in mind the previous point: from the fact that the actor gains satisfaction from the spiteful action it doesn't follow that gaining satisfaction is the ultimate motive.) Psychological egoism cannot accommodate truly spiteful motivations.

(Psychological egoism must not be confused with *ethical* egoism, which is a normative theory claiming that all human actions *ought* to be motivated by selfish motives. Ethical egoism is often presented in a confused manner and is arguably not even coherent (Medlin 1957; Gauthier 1974); but here it is identified only to avoid confusion.)

Standing in stark contrast to psychological altruism is evolutionary altruism. Whereas "altruistic" in the psychological sense applies to actions done from a certain kind of motive

(and can also denote this motive itself), "altruistic" in the evolutionary sense applies to *traits*. A trait is evolutionarily altruistic if and only if it benefits another at some cost to the individual, where benefits and costs are understood in terms of reproductive fitness. It must be added that the trait has been selected for *because* it benefits another, otherwise one ends up counting as altruistic such things as the kiwi's flightlessness which renders it such easy prey for cats.

In characterizing evolutionary altruism we face a theoretical decision over whether to consider reproductive fitness over some restricted time period or over the life of the individual. When an organism provides costly help to another, and this trait has evolved because the organism reliably receives a proportionally greater pay-off at a later time, then this is widely referred to as "reciprocal altruism" (Trivers 1971) (*see* RECIPROCITY). Consider, for example, grooming among various primate species, where an important factor determining whether individual A grooms B is whether B has groomed A on an earlier occasion (Schino and Filippo 2010). B's grooming A on that earlier occasion may look like an instance of evolved costly helping and thus might be classed as evolutionary altruism. But if we are taking into account the reproductive costs and benefits to B over the long-term (e.g., over its life-time), then we must include the countervailing benefits that B gains from being groomed by A, in which case B's grooming behavior is not an instance of evolutionary altruism at all, since in performing this behavior B is reproductively better off (eventually) than if it did not. (See West et al. 2007.) (Nor does A's reciprocal grooming count as altruistic, since we may assume that this behavior will be part of an ongoing relationship with B, and thus A is better off reciprocating.) For this reason many theorists renounce the label "reciprocal altruism" in favor of the less confusing "reciprocity."

Another theoretical decision to be made is whether to consider the reproductive costs and benefits *inclusively*—that is, whether to consider only those costs and benefits that are achieved via direct reproduction, or whether also to include costs and benefits that are achieved indirectly—say, by forfeiting one's own reproductive possibilities in order to help siblings procreate. When an organism provides costly help to another, and this trait evolved because the recipient is a close kin member, this is widely referred to as an instance of "evolutionary altruism." Consider, for example, male wild turkeys helping brothers in courtship displays (Krakauer 2005). The helper incurs costs of time and energy which detract from his own breeding possibilities, while directly advancing those of another, and thus this behavior may look like a textbook case of evolutionary altruism. Again, however, if we are counting fitness benefits inclusively then the organism performing this helpful behavior may be better off than if it did not—since, under certain conditions, by ensuring that its brother breeds it does better (regarding replicating its genes in future generations) than by putting its time and energy into a failed attempt at breeding itself—and thus the behavior is not an instance of evolutionary altruism at all.

References to "altruism in nature" are widespread: the term is used in regard to bees, ants, food-sharing in primates, pack-hunting in canines, warning calls in rodents, and so forth. Talk of *psychological* altruism in such cases is nearly always entirely misplaced: the organisms in question do not satisfy the psychological prerequisites for having ultimate motives concerning others' or their own welfare. They do not, for example, have the concepts of *other* and *self* that are necessary. (The only uncontroversial satisfiers of these prerequisites are humans.) Aware of this fact, those speaking of "altruism in nature" are usually quick to insist that they refer only to the *evolutionary* variety. But the above considerations show that much

confusion often remains. If, for example, the suicidal sting of the bee is explained by reference to kin selection, then rather than this being an explanation of how an altruistic trait evolved, it may be considered an explanation for why it is not really altruistic at all. It is, in fact, far from obvious how much evolutionary altruism exists in nature which cries out for explanation. What we certainly do observe in nature is a great deal of *cooperative* behavior (some of which *appears* to be evolutionarily altruistic), and kin selection and reciprocity are probably important processes in explaining *that*.

One process that could produce true evolutionary altruism is group selection. (See Sober 1988; Sober & Wilson 1998.) Suppose individual A has a trait of providing help to others in a way that renders its own (inclusive) reproductive fitness worse off than that of the beneficiaries, and A receives no recompense for its troubles. Suppose, however, that the *group*—containing A and beneficiaries—does better in reproductive terms than competing groups (which do not contain helpful individuals like A). In such a situation, A may have a lower reproductive fitness than members of its own group, but a higher fitness than members of the global population consisting of the ensemble of competing groups. Here there would be selective pressures pulling in both directions: *within the group* there will be selection against such a self-sacrificial helpful trait; *between groups* there will be a countervailing pressure in favor of its emergence. With the appropriate population structure and the right balance of costs and benefits, the between-group selective pressure could be stronger, producing genuine evolutionary altruism. There is much dispute, however, about how frequently this occurs in nature. (See Maynard Smith 1964; Okasha 2001; Wild et al. 2009.)

Without pausing to investigate the details of how much cooperation in nature really is evolutionarily altruistic, one can at least safely say that cooperation often turns out to be evolutionarily selfish, in the sense that the cooperative behavior ultimately enhances the actor's reproductive fitness better than not cooperating. The temptation that it is crucial to resist is thinking that this evolutionary selfishness has any bearing on psychological selfishness. The comment above about psychological prerequisites is sufficient to demonstrate this. The helpful trait in question may belong to a plant (spreading its roots less aggressively when its neighbor is a sibling plant, say), and from the revelation that this is an evolutionarily selfish trait it would obviously be ridiculous to conclude that it is a psychologically selfish trait. A plant cannot be psychologically selfish or altruistic; it cannot be psychologically anything.

Though obvious when applied to plants, the distinction frequently goes missing when applied to humans, leading to much confusion. It is plausible to assume that humans have evolved various mechanisms governing cooperative interactions, and it is possible that these adaptations are really evolutionarily selfish. But it wouldn't follow that the cooperative behaviors in question are therefore in any sense "really" psychologically selfish. It would be mistaken to assume, for example, that if humans have evolved mechanisms overseeing reciprocal exchanges, then these exchanges must be motivated by the promise of future payback. Of course we all know from experience that human reciprocal relations are *often* motivated by selfish expectations, but they need not always be. Perhaps the most effective way for natural selection to encourage stable human reciprocity was to create psychological tendencies to care directly for the welfare of one's exchange partners. In short, the fact that interactants repay the favor may explain how the trait evolved without explaining why any agent acts. (See Tinbergen 1963.)

Indeed, it is not implausible that natural selection has forged humans to have psychologically altruistic tendencies. A parent motivated to aid his distressed child simply because he loves the child is—one might plausibly claim—moved by a more reliable and less complicated process than a parent moved via a combination of the belief that the child's suffering has a negative effect on his own welfare plus his love for himself. (See Sober 2000.) By analogy, a person prompted to pull her fingers from a flame *by pain* seems moved by a more reliable and less complicated process than a person who forms a belief about the bodily damage caused by fire and calculates the costs and benefits of action versus non-action. This argument may not be without problems (see Stich 2007), but whether successful or not it serves to emphasize the crucial point: Even if human psychology was forged in accordance with principles of natural selection, in ways that increased the reproductive fitness of our ancestors (i.e., producing evolutionarily selfish traits), the mechanisms themselves may be dispositions for genuine psychological altruism.

What bearing do any of the above observations have on the topic of human *morality*? There may well be subtle connections, but it is important at the outset to acknowledge that none of them are obvious. For a start, let us put *evolutionary* altruism to one side. A plant may have evolutionarily altruistic traits, but the plant neither makes moral judgments nor is a suitable subject of our moral appraisals. Similarly, from the fact that a person acts from moral motives in a manner that earns our moral praise, one can conclude nothing about whether those actions enhance another's fitness at reproductive cost to the actor.

That there are connections between *psychological* altruism and morality is slightly more plausible, but one must first observe the disconnections. In order to be psychologically altruistic a creature needs to be fairly cognitively sophisticated (so plants are certainly disqualified), but it doesn't follow that the creature is therefore capable of making moral judgments. This truism is potentially muddied by the fact that the only clear-cut case of a species capable of psychological altruism (and selfishness) is also the only clear-cut case of a species capable of making moral judgments: namely, humans. Still, the conceptual distinction does not seem difficult to discern. One can imagine members of a cognitively sophisticated social species, motivated by love and altruistic tendencies towards their fellows, but who fail to "moralize" these feelings—who are, in fact, constitutionally incapable of making a moral judgment. Such creatures have powerful desires to see their loved ones flourish, but cannot conceive of actions satisfying those desires as morally right or obligatory. It might be acknowledged that these imaginary creatures don't make moral judgments but maintained that they are at least morally *praiseworthy* (that is, that they warrant *our* moral judgment). But upon reflection even this is unclear. After all, altruistic motives can prompt someone to act in a morally despicable manner. Consider a mother who genuinely adores her child, who wouldn't hesitate sacrifice her own interests for the child, and who poisons all the other children at the sports day so her child can win.

The fact that a capacity for altruistic motivation does not imply any moralizing capacity is important to note when seeking evidence concerning the "evolution of morality." Suppose one is investigating the thesis that the human capacity to make moral judgments is a biological adaptation—a view known as "moral nativism." (*See* EVOLUTION, ETHICS AND.) Evidence that certain human traits are evolutionarily altruistic would have no obvious bearing on the nativist thesis. Nor would evidence that humans are by nature cooperative be sufficient to establish the thesis. Nor would mustering evidence that natural selection has forged humans to have the capacity to love and care directly for the welfare of each other

("psychological altruism nativism") help to establish moral nativism. For example, the argument mentioned above—that human psychological altruism may have evolved because it is more reliable and simpler than alternative psychological mechanisms—is not an argument for moral nativism.

Nevertheless, an argument for moral nativism with a matching structure might be developed. It might be claimed that for certain kinds of social interaction a psychological mechanism that governs compliance via establishing thoughts of *obligation*, *wrongness*, *virtue*, etc., is more effective (and thus, ceteris paribus, more adaptive) than other possible mechanisms. For example, an individual, A, who conceives of some cooperative action toward B as *obligatory*, say, may have a more stalwart motivation to perform that action than someone who thinks of that action as a prudentially good idea. (See Joyce 2001 chapter 8; Joyce 2006 chapter 4.) Someone moved by such moral motivations may enjoy adaptive advantage even over someone who cooperates because she really loves B and cares altruistically for B's welfare. To see why this may be so, note how failure to comply with a moral norm licenses thoughts of "punishment being deserved," in a way that failure to act prudently or failure to benefit a loved one do not license. Such thoughts of punishment and desert may play an important role in combatting various kinds of motivational failure, such as weakness of will or discounting the future when calculating costs and benefits.

It is easy enough to construct a plausible hypothesis according to which moral thinking was adaptive to our ancestors, and on that basis speculate that it is an evolved human trait. However, locating concrete evidence in support of the position (or, indeed, against it) is a far more complicated and contentious program. Opposing hypotheses—that moral thinking is a fairly recent cultural artifact, for example, or that moral thinking is a byproduct of other evolved psychological traits—are not the least far-fetched. (See Nichols 2005; Sterelny 2010.) One factor complicating the debate is that there is no single hypothesis deserving the name "moral nativism"—the above characterization leaves much open. At one end of the spectrum is the view that particular moral judgments are a design feature; at the other end is the view that the content of morality is acquired from the social environment but that there is an innate mechanism dedicated to this kind of learning.

Having noted that evidence for psychological altruism is not evidence for moral judgment, is there anything to be said for the reverse implication? Might moral judgment imply altruism? First we need to distinguish between the appraiser and the appraised: Let us stipulate that Fred makes a positive moral judgment concerning Sally's action. Altruistic motivation appears to be necessary for neither party. Sally (the appraised) need not be acting altruistically; she may be moved by a strong sense of duty (to keep a promise, say), with desires for the welfare of others playing no substantive role in her motivations. That altruism is unnecessary for the appraiser (Fred) is even more obvious. After all, he is not performing any *action* at all, but simply making a judgment, so it is not clear to whose welfare we should imagine any altruistic motivation on his part being directed. (The answers "Sally's" or "his own" seem equally puzzling.)

Perhaps the only implication that has a glimmer of plausibility is the claim that in order for Sally to be morally judged she must at least have the *capacity* for altruistic action, even if her current action is not so motivated. But whatever plausibility this may have evaporates when the distinction is observed between what it takes for someone to be the subject of an *appropriate* or *warranted* moral judgment, as opposed to being the subject of a moral judgment *per se*. Suppose Sally in fact has no capacity for altruistic motivation at all; perhaps

a childhood injury left some crucial part of her brain impaired; but apart from this she is normal and an upstanding citizen for self-regarding reasons (because, let's say, she wants to get into heaven). One might think that in such circumstances it would be *incorrect* for Fred to judge one of Sally's actions in a positive moral light—especially if we credit him with knowledge of her condition—but that is beside the point. To support the claim that the capacity for altruism in the appraised is necessary for moral judgment one would need to maintain that in these circumstances it is *impossible* for Fred to morally judge Sally positively. And this seems highly implausible.

Both psychological altruism and moral judgment are biological phenomena that draw on evolved psychological capacities; the nativist question concerns whether there are faculties evolved specifically to perform these tasks, or whether we are simply using more general psychological mechanisms that evolved for other purposes. What has been stressed here is the need to tease the two phenomena apart. Psychological altruism is certainly an important part of morality—perhaps its cultivation is a central moral value—but it is neither necessary nor sufficient for either moral judgment or moral praiseworthiness.

CROSS-REFERENCES:

Egoism; Evolution, Ethics and; Reciprocity

REFERENCES:

Feinberg, J. 2008. "Psychological egoism." In R. Shafer-Landau & J. Feinberg (eds.), *Reason and Responsibility*. Thirteenth edition. Belmont, CA.: Thomson Wadsworth, pp. 520-532.

Gauthier, D. 1974, "The impossibility of rational egoism." *Journal of Philosophy* 71, pp. 439-455.

Joyce, R. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.

Joyce, R. 2006. *The Evolution of Morality*. Cambridge, MA.: MIT Press.

Krakauer, A. 2005. "Kin selection and cooperative courtship in wild turkeys." *Nature* 434, pp. 69-72.

Maynard Smith, J. 1964. "Group selection and kin selection." *Nature* 201, pp. 1145-1147.

Medlin, B. 1957. "Ultimate principles and ethical egoism." *Australasian Journal of Philosophy* 35, pp. 111-118.

Nichols, S. 2005. "Innateness and moral psychology." In P. Carruthers, S. Laurence, & S. Stich (eds.), *The Innate Mind: Structure and Contents*. New York: Oxford University Press, pp. 353-369.

Okasha, S. 2001. "Why won't the group selection controversy go away?" *British Journal for the Philosophy of Science* 52, pp. 25-50.

Schino, G. & Aureli, F. 2010. "The relative roles of kinship and reciprocity in explaining primate altruism." *Ecology Letters* 13, pp. 45–50.

Sober, E. 1988. "What is evolutionary altruism?" In M. Matthen & B. Linsky (eds.), *New Essays on Philosophy and Biology*. (*Canadian Journal of Philosophy* suppl. vol. 14), pp. 75-99.

Sober, E. 2000. “Psychological egoism.” In H. LaFollette (ed.), *The Blackwell Guide to Ethical Theory*. Oxford: Blackwell, pp. 129-148.
Sober, E. & Wilson, D. S. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
Sterelny, K. 2010. “Moral nativism: A sceptical response.” *Mind & Language* 25, pp. 279-297.
Stich, S. 2007. “Evolution, altruism and cognitive architecture: A critique of Sober and Wilson’s argument for psychological altruism.” *Biology and Philosophy* 22, pp. 267-281.
Tinbergen, N. 1963. “On the aims and methods of ethology.” *Zeitschrift für Tierpsychologie* 20, pp. 410-433.
Trivers, R. L. 1971. “The evolution of reciprocal altruism.” *Quarterly Review of Biology* 46, pp. 35-57.
West, S. A., Griffin, A. S. & Gardner, A. 2007. “Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection.” *European Society for Evolutionary Biology* 20, pp. 415-432.
Wild, G., Gardner, A., & West, S. 2009. “Adaptation and the evolution of parasite virulence in a connected world.” *Nature* 459, pp. 983-986.

SUGGESTED READING:

Alexander, R. 1987. *The Biology of Moral Systems*. NY: Aldine de Gruyter.
Batson, C. 2011. *Altruism in Humans*. Oxford: Oxford University Press.
de Waal, B. M. 2006. *Primates and Philosophers*. Princeton, NJ.: Princeton University Press.
Fehr, E. & Fischbacher, U. 2003. “The nature of human altruism.” *Nature* 425, pp. 785-791.
James, S. 2011. *An Introduction to Evolutionary Ethics*. Oxford: Wiley-Blackwell.
Katz, L. (ed.). 2000. *Evolutionary Origins of Morality*. Exeter, UK: Imprint Academic.
Kitcher, P. 1994. “Four ways of ‘biologizing’ ethics.” In E. Sober (ed.) *Conceptual Issues in Evolutionary Biology*. Cambridge, MA.: Bradford Books, pp. 439-452.
Mansbridge, J. (ed.). *Beyond Self-interest*. Chicago: University of Chicago Press.
Milo, R. (ed.). 1973. *Egoism and Altruism*. Belmont, CA.: Wadsworth.
Monroe, K. R. 1996. *The Heart of Altruism*. Princeton, NJ.: Princeton University Press.
Oliner, P. M. et al. (eds.). 1992. *Embracing the Other*. New York: New York University Press.
Wilson, D. S. 1991, “On the relationship between evolutionary and psychological definitions of altruism and egoism.” *Biology and Philosophy* 7, pp. 61-68.