

ep

Richard D. Alexander, manuscript

Essay on Self-Interest: RDA, 12 March 2007

SELF-INTEREST AND ITS ALTERNATIVES

Unsettling or not, there are reasons for being skeptical that beneficent acts accompanied by pleasure to the benefactor can be accepted uncritically as alternatives to self-interest.

Some biologists, economists, philosophers, and social scientists have recently gone astray with respect to the topic of self-interest. This essay is intended to give some possible reasons:

Darwin said that natural selection cannot evolve a trait that causes "hurt" to its bearer.

To understand this statement we need to explain two concepts: (a) evolved trait and (b) "hurt."

Darwin also said that if anyone could find a complex organ that could not have evolved through numerous, successive, slight changes, he could "absolutely destroy my theory." The numerous successive, slight changes to which Darwin referred we now know reflect the fact that gene mutations – especially beneficial ones – generally produce only "slight changes." As Fisher pointed out, the larger the effect of a gene mutation, the less likely it is to be (or later become) beneficial. Massive changes that are random with respect to whether or not they help the organism (species) in which they occur are extremely unlikely to help that organism, partly because they are likely to be so detrimental that they disappear quickly and have no chance to be beneficial later because of other changes in later generations. It is also true that organisms have become so extremely complex that nearly all changes that are random with respect to whether they benefit the organism, whatever the size of their effect, are bound to be deleterious.

An enormous accumulation of studies of genetic variations in organisms and their effects supports Darwin's idea. Included is the fact that differences between species are also owing to numerous, successive, slight changes, denying creationists' arguments that species differences cannot be explained as a result of small step-by-step evolutionary changes. This fact has been demonstrated by large numbers of interspecific hybridizations, which invariably show that between-species differences are of the same nature – accumulation of numerous, successive, slight changes -- as between-individual differences within species. Macro-evolution really is micro-evolution extended.

An evolved trait, then, is one that is built upon a history of Darwin's numerous, successive, slight changes. This definition excludes single changes, such as individual mutations, accidents (such as learning gone awry), and evolutionary novelties resulting from rapid environmental change (such as newly experienced behavior-altering drugs).

What is a "hurt" to an organism? Organisms perpetuate themselves only in the sense that they transmit their genes to the next generation in ways that enable those genes to re-create individual organisms that can do the same. To "hurt" an organism, then, is to limit, restrict, or prevent its reproductive success.

So Darwin was making the commonsense and apparently irrefutable statement that differential reproduction cannot result in step-by-step changes in an organism if those successive steps of change

with my ~~notes~~ attached.
response

“hurt,” or represent an accumulation of net costs to, the organism’s reproductive success. This means that net-cost acts by any organism, including humans, will by definition lower the likelihood of their own perpetuation, or the reproduction of whatever it is that causes them to be net-cost acts. There is no reason to believe that net-cost social or altruistic acts should be excluded.

Pleasure and pain are concepts that can be used to speak generally about positive and negative responses to actions, and positive and negative reproductive effects of actions. Pleasure and its relatives typically reflect actions or effects that cause the pleased individual to seek out and repeat the act. Conversely, pain typically causes an individual to withdraw and avoid repeating the act that caused the pain. Pleasure and pain, like all other traits, presumably evolve by Darwin’s numerous, slight, successive changes. Unless the definitions just given are wrong, it is highly unlikely that an act can evolve to cause pain unless its bearer gains reproductively by foregoing or avoiding that act. It is also unlikely that any act can evolve to cause pleasure unless it improves the reproductive success of the organism deriving the pleasure. There does not seem to be any reason to exclude from these considerations acts of beneficence toward others. If such acts consistently result in net costs they will evolve to cause pain, in one or another of its versions, such as disgust, withdrawal, or avoidance. If such acts result in pleasure or satisfaction, they will evolve to cause repetition of the act.

When individuals pass benefits to close relatives such as offspring or siblings, they typically experience pleasure, even if they have little or no expectation of the offspring or siblings ever returning any benefits at all to the phenotype of the beneficent relative. This act is, however, consistent with self-interest because, if our lifetimes are shaped by evolution, then our very existence means that self-interest has to be defined as including effects on genes. Inescapably, self is both phenotype and genotype. A lifetime is (reproductively) successful only if it results in the genes of the one doing the living being transmitted to the next generation in such fashion as to cause the transmission to continue across successive generations. William D. Hamilton noted that we should expect organisms to evolve to treat their genetic relatives according to their degree of relatedness, when such qualifiers as amounts of assistance from other sources and likelihood of turning benefits into reproductive success are controlled. This description is consistent in every respect with what has been said so far about how evolution proceeds, and how pleasure and pain evolve. Attending to one’s close relatives cannot, therefore, be automatically viewed as an alternative to self-interest. Instead it is the central theme of self-interest.

We know that individual humans constantly do things that benefit others, including non-relatives, and that they often experience pleasure as a result. We also know that there are many different ways by which individuals that show kindness or other benefits to others can be rewarded. Unless beneficent acts are consistently accidental or novel, it behooves us to seek out all possible returns to beneficent individuals that experience pleasure from their kind acts to see if, as predicted, they are likely to receive return benefits that cause the entire set of acts to benefit them. If we have good reason to doubt that a particular socially beneficent act results in such over-compensating returns, then we should expect to find evidence of negative rather than positive reactions in the originally beneficent individual. Conversely, a pleasurable effect in the extending of beneficence necessarily causes us to assume that there is a significant likelihood of sufficient returns to make the beneficent act a net benefit to the giver as well as the receiver.

We also know that many acts of apparent beneficence are consciously intended to result in over-compensating returns. Thus, when a person deposits money in a bank, he or she consciously expects the banker to return more money than was deposited. The depositor also realizes, consciously, that the banker will gain from use of the deposited money. Both parties thus expect to gain, and they both know it throughout the entire transaction. This is an entirely conscious instance of social reciprocity. No alternatives to self-interest are involved, and presumably both parties experience pleasure from the exchanges involved. This interaction is therefore consistent in every respect with Darwinian evolution. We know this only because all of the interaction is conscious, as is often the case with the experiments of economists involving monetary exchanges.

The difficult question is how to interpret cases of beneficence in which the second "half" of the potentially reciprocal act is apparently not conscious (excepting the pleasurable feeling from having carried out the beneficence) and may take one or more of many possible forms. These are beneficent acts for which we cannot immediately identify sources of overcompensating returns, and may never be able to do so consciously. If the act causes pleasure in the beneficent individual – and we expect to generate an accurate explanation of the beneficent act – then, before drawing conclusions about self interest or its alternatives, we are obligated to search throughout the entire array of possible returns for likely sources of overcompensation for the beneficence. At this time we probably are not even close to being consciously aware of all such possible returns, or of the particular returns most likely in specific cases; it is unlikely that we ever will be unless we make specific, intensive, extended searches for all such returns. This is a complex assignment, which demands extraordinary insight into the details of everyday human sociality.

Here is part of my imperfect and incomplete discussion of possible returns from social beneficence in: *The Challenge of Human Social Behavior: A Review Essay Stimulated by Hammerstein, Peter (ed). Genetic and Cultural Evolution of Cooperation (MIT Press). Evol. Psychol. 4(2):1-28 (2006):*

In my 1987 [The Biology of Moral Systems] discussion of indirect reciprocity (the most recent and best source [at that time] for a review of my views on that topic), I said the following: "Returns from indirect reciprocity may take at least three major forms: (1) the beneficent individual may later be engaged in profitable reciprocal interactions by individuals who have observed his behavior in directly reciprocal interactions and judged him to be a potentially rewarding interactant (the 'reputation' or 'status' of the investor is enhanced, to his or her ultimate benefit), (2) the beneficent individual may be rewarded with direct compensation from all or part of the group (such as with money or a medal or social elevation as a hero), and (3) the beneficent individual may be rewarded by simply having the success of the group within which he or she behaved beneficently contribute to the success of his or her own descendants and collateral relatives."

Leaving aside maladaptive accidents or errors, I can think of at least three other possible adaptive (indirect reciprocity) explanations for what some authors in this volume call "one-shot" social investments: (4) the return may be to the beneficent individual's relatives or friends, and the nature of social information spread as a result may be such as to make this kind of return consistent with investment being self-serving; (5) The investment may also be done as part of the individual practicing how to engage in reciprocity adaptively, as with individuals who practice while alone for success in, say, being humorous, or in lecturing, or in developing a useful

conscience, or any other social behavior (i.e., as a way of learning how to invest socially in a more rewarding – more profitable – way). This is not reciprocity per se, nor is it evidence of net-cost beneficence; rather, it is investment that in the end functions to improve the individual's later engagement in social reciprocity. In my experience we do this kind of thing all the time, and I regard it as an essential part of knowing how to behave socially in one's own interests [it can happen during "one-shot" experiments, with the results of the practice having its rewards via interactions with others (or information passed to others) following the experiments, and even planned that way by the participants during the experiment. In an important sense there may be no two-party social interactions in humans.]. Finally, (6), the investment may be part of an individual's effort explicitly to elevate the general level of reciprocity in society. Thus, generous donations to people affected by a disaster, or efforts to enlist in the armed forces at the onset of war, can have snowballing effects on donations or rates of enlistment that raise the level of social investment generally – and may benefit the individual by various indirect reciprocity returns. Whenever the general level of social investment is raised, all persons had better pay attention and act wisely, or they may suffer from being viewed as laggards, or unduly self-serving and stingy. It is not necessary, however, that the reputation of any individual contributing to the rise in social investment be involved. Such donations can be entirely anonymous and the donor can still gain, as whenever his or her interests are close to those of the entire group.

Any experiment done in a simplified laboratory situation must be monitored with extreme care to prevent the results from being misinterpreted because of responses of the participants to at least the more obvious ways in which the experiment does not reflect "real-life" situations. For example, in some everyday social situations, return beneficence may occur years after an act of beneficence, and may involve long and complex chains of beneficent acts and long and complex chains of participants – and may on those accounts serve the original benefactor better. Conscious understanding of such effects, moreover, may remain imperfect or absent. It is also possible that beneficent acts are sometimes triggered by non-conscious cryptic signs of probable return beneficence to the originally beneficent individual, or that special aspects of a beneficent individual's behavior with respect to having carried out the act of beneficence may trigger eventual social returns. In any case, it is reasonable to interpret beneficence accompanied by pleasurable effects as a social investment with either conscious or unconscious expectation of return rather than as net-cost altruism.