

The Evolution of Commitment and the Origins of Religion

Full version of an article that appeared in
Science and Spirit 10 (2), July/August 1999, p. 32-33, 46

Randolph M. Nesse
The University of Michigan
5057 ISR
Ann Arbor, MI 48106-1248

I once talked with a woman who was full of hateful feelings for her mother, but was sure that her mother loved her. “How are you sure?” I asked. “Because,” she replied, “she showed me the book where she wrote a note every time she did something nice for me.”

The very idea of keeping social accounts with our kin and closest friends is repugnant. Our generous gestures are supposed to come from the heart. We help our friends because we care about them, not because we expect their help in return. We count on their help when we need it, not when we are in a position to repay them later. It is fine to make exchanges explicit in business, but in close relationships, it is taboo. When evolutionary biologists explain our closest relationships with non-relatives as trading favors, most people dismiss them as cold-hearted cynics, or point the differences between humans and animals.

But there is a real mystery here. How is altruism possible? Certainly we get benefits from trading favors. And helping other related individuals benefits your genes, so that is easy to explain. But what about deep friendships, where people help each other even, especially, when there is nothing to be gained? And what about human behaviors that arise from moral commitments? Must the capacities for such behaviors somehow give a selective benefit?

The surprising and disturbing answer is almost certainly, “Yes.” While each particular action need not help our genes, our capacities for personal feelings of commitment and moral obligation could have been shaped by natural selection only if they somehow give an advantage to the genes that make them possible. This is the crux of what seems to me to be a central moral problem of our time. Prior to the 1960s, most biologists thought that organisms evolved to act in ways that benefited the group. But then three biologists changed all that. At about the same time, George Williams showed that group selection was feeble, and William Hamilton discovered an alternative explanation for altruism—actions that decrease an individual’s reproduction can nonetheless evolve if they provided enough benefit to relatives who have the same genes. Then, in 1971, Robert Trivers wrote his classic paper on reciprocal altruism, which revealed that many good deeds have a pay-off in favors returned. The old view crumbled, and the resulting tumult continues. In the past decade, for instance, more than a score of books have appeared about evolution and morality. Something big is happening to the way we understand ourselves, and it does not seem to be nearly settled.

In the midst of all this, several people have noted the adaptive value of social strategies based on “commitment.” The economist

Jack Hirshleifer was an early proponent of the idea, which was brought to many people's attention in Robert Frank's book, *Passions Within Reason*. Earlier, Thomas Schelling had described the crucial role of commitment in political strategies, including nuclear confrontation. By making a commitment to act, in some future situation, in a way that would not then be self-interested, an individual or group can change the expectations, and thus the behaviors, of others.

A commitment to future generosity, for instance, can make possible relationships far deeper than those based on mere reciprocity. People who have capacity to engage in such relationships have a huge advantage. Instead of getting favors only when they can return them, they can realistically hope for help from friends when they need it the most, when they are sick, alone, or poor. Of course, to get these benefits, they must be ready to give them as well. This is favor trading of a special sort. Instead of limiting relationships to the giving and receiving of favors, some people make commitments that may require action that is not in their best interests. If required, they usually fulfill those commitments. The social contract is based not only on trading favors when there is expectation that they will be returned, but also on trading commitments to act, in the future, in ways that would not then be beneficial.

Not all commitments are benevolent, however. A commitment to spiteful retaliation can be equally useful. As Frank notes, a person whose briefcase was stolen might spend extra expensive days in a distant city just to testify against the thief. Such spite is, from a rational point of view, senseless. Until, that is, you consider reputation. If a potential thief knows a particular person is committed to spiteful retaliation, he won't touch that person's briefcase. On a larger, more stark and terrifying level, if one country knows that an attack on another country will be met with a nuclear exchange that would largely

extinguish human life in the hemisphere, it will be deterred from the attack. This is the basis for the nuclear strategy of mutual assured destruction.

Commitment strategies generate extraordinary paradoxes. What is the best way to convince others that you really will engage in spiteful retaliation? By creating a system in which the retaliation is automatic and out of your control, or by acting erratic and emotional. As Frank has pointed out, certain emotions seem to fulfil these functions nicely. The long history of debate on whether emotions are opposed to reason or in the service of reason, may be clarified by this fine point of game theory. To be effective, a signal of commitment must be accompanied by demonstrations of irrationality, otherwise, others will conclude that the person will, in the end, do what would give the maximum rationally calculated gain. There are people who always do what is, in this sense, rational, especially those with the rationalistic, duty-bound, characteristics of obsessive compulsive personality disorder (a problem quite different from the anxious preoccupation with preventing harm to others that is typical in obsessive compulsive disorder). They are at huge disadvantage because they can't use commitment strategies effectively. Erratic, emotional, and irrational tendencies are essential signals that natural selection may well have shaped to make commitment strategies possible.

Examples of such phenomena are everywhere. A couple who only a few years ago were deeply in love, now are divorcing and trying to hurt each other in every way possible. Each tells their friends about all the awful things the other has done, thus exposing their own sins as well. They know it is spiteful, but their anger propels them. On the more positive side, a spouse may devote his or her life to taking care of a dying partner, a commitment that may take up some of the best years of life with no tangible payback possible.

In short, people may have been shaped by natural selection to have a capacity for passionate commitments. We may have the capacity to behave according to emotions and principles, instead of calculation, because strategies based on commitment give such a big advantage. In one sense these emotions are primitive, but on the other hand they make possible social strategies vastly more complex than those in most animals. A commitment that is rational, in the sense that a pay back is expected, is mere reciprocity. You can play tit-for-tat and do just fine. When we seek partners in life, however, we don't want someone who will be there when we have something to offer, we want someone who will be there when there is nothing we can offer. So, when we detect that supposed friends are acting on the basis of calculation instead of feeling, we wonder if they will really be there for us when the going gets tough. And then, we test them. This testing, the ethologist Zahavi called it "testing of the bond," is really important. Lover's spats give a chance to see if your potential partner can put up with some hassles and mistreatment. If not, it is better that you know now, so you can look for a partner who is more trustworthy. Even birds seems to do this. A male cardinal, when wooing his potential mate, must put up with all kind of abuse as he bring his gifts of seed. Only after extensive demonstrations of such devotion, will the female agree to mate.

Another paradox of commitment strategies is that what a person believes, changes his or her actual social world. If a person believes that other people are selfish and untrustworthy, then that expectation will be correct for that person and makes commitments senseless. A person who cannot make emotional commitments is sure not to get any in return. For that person, the world really is a cold and heartless place. On the other hand, people who believe that love is possible, based on their earliest experiences in life, are able to form relationships based on emotion,

relationships far stronger than those based on exchange. They are also susceptible to manipulation and terrible disappointment. The capacity for "basic trust" may be a social construction built on mental foundations shaped by natural selection. It may develop early in life, or whither, depending on the child's experiences. Those of us who do psychotherapy know the extraordinary difficulty of trying to help a person gain such an ability to trust after many experiences of betrayal.

I suspect that this is why we humans are so concerned with our views of human nature. What we believe about human nature has tangible effects not only our lives, but also on the structure of our societies. While it is perfectly clear to me that humans were shaped by natural selection, I share the fears of those who worry that interpreting human behavior in crude terms can make social life even worse than it often is. If people learn that all apparent altruism is mere exchange or worse, manipulation, then their capacity to believe in love is corroded. And if they cannot believe, then for them, love is impossible.

I wonder if this might be one of the main reasons that so many people deeply value religion. It is difficult to create committed relationships, one by one. You never know if the other person will live up to the commitment. But if you are a member of a group, and if everyone in the group makes sacred vows to follow certain rules, especially rules to help each other when there is no hope of reward, and if they monitor each other to be sure that all are following the rules, this can create a community of believers. This may explain why it is so important that belief be based, as Kierkegaard emphasized, on faith itself, and not on reasons. Communities of believers are networks of fictive kin that can provide huge benefits for their members. This may explain why religious fervor is responsible for so much good in the world. Paradoxically, this same capacity for subjective commitment may explain why

religions have also been responsible for so much evil, whether in the form of crusades against out-group members, or drastic enforcement of conformity within the ranks.

There are inherent problems and paradoxes in commitment strategies. First, it is always tempting to promise more than can be delivered. Some excess solidifies the commitment, but more arouses skepticism and testing. Second, so many benefits come from having committed friends, that people are eager to create such relationships. Soon, there are too many to make good on each commitment, or one finds oneself in a triangle, committed to helping two people who are fighting. Some commitments, especially marriage and political alliances, are defined, in part, by prohibition of other commitments. And then, of course, there are sociopaths who promise everything, and deliver only enough to deceive you, then

depart with whatever you gave them, and your sense of trust. This inevitable dance of deception, and self-deception in social relationships may have shaped its own mental mechanisms. Not only do others deceive us, but we often find that our own motives are not as pure as we would like to think. And we may have tendencies to stay completely unaware of some betrayals by others. People who are consciously aware of most of their own motives may have a very hard time of life. Compared to social life, chess seems simple. Commitment strategies give rise to complexities that may be a selective force that has shaped human intelligence. Perhaps this is why human psychology and relationships are so hard to understand. Perhaps a better understanding of commitment will illuminate the relationships between reason and emotion, and biology and belief.