

Running head: ORIGINS OF HUMAN MORALITY

The Origins and Implications of Human Morality

Cameron Clark

Simon Fraser University

## Abstract

Firstly, a case is made for the development of human morality by the process of evolution by natural selection. Cooperation and reciprocal altruism are examined in the context of physiological and cultural development over time. Secondly, it is shown that the misunderstanding of this process leads to the fear of “moral imperfectibility” via the is-ought distinction. Finally, a case for genuine moral progress is made by appealing to the continual increase of participation in non-zero-sum cooperative ventures throughout human history. In short, a case is made against the notion that a naturalistic (i.e. biological) explanation of human morality is to be feared and/or avoided.

### The Origins and Implications of Human Morality

We all have an intuitive sense about what is right, or at least what we “ought” to do, and consequently, what we “ought not” to do. The majority of people would consider it wrong to cheat, murder or steal, although would have a difficult time articulating how these moral suasions came about. However, as Kohlberg showed, there is no short supply of reasons that people might offer as to why they act the way that they do when confronted with a moral dilemma. People may justify their actions by appealing to such diverse motivations as avoiding punishment, adherence to a law or social contract, or even adherence to self-decided principles of right and wrong (as cited in Krebs, 1998). But where do these motivations come from? Are they innate, or are they learned through social and cultural processes? More importantly, what implications does each of these views hold for the future of human morality?

Evolutionary theorists make the case for an innate moral sense forged over tens of thousands of years by the selection pressures of our evolutionary environment. It is thought that morality, like language, is a powerful tool that can be used to increase the overall adaptedness of an animal, and thereby confer greater reproductive success on those who possess it (Trivers, 1971). On this view, those that possessed moral qualities such as trustworthiness and a hard work ethic would have a greater chance of surviving and reproducing than those that do not. However, other thinkers doubt the explanatory power of the evolutionary view, and may even fear that such a biological understanding of mind and morality will lead to a total collapse of traditional moral values. For example, if violence and selfishness are proven to be innate, is it not the case that we are forever doomed to moral imperfection (Pinker, 2002)? Better that we do not open Pandora’s Box say the proponents of this view; if we already know that murder and rape and outright selfishness are wrong and unwanted in our society, why bother trying to condone or dignify these

actions with an evolutionary theory? It is the purpose of this paper to demonstrate that human morality is perfectly explicable as a product of evolutionary processes. Secondly, I will briefly address the fear of moral imperfection associated with this view of human morality. Finally, I falsify these fears by demonstrating that moral progress can continue to occur over time, given the fixed moral mechanisms endowed to us by evolutionary processes. In short, I will make a case against the idea that a biological understanding of human morality is something to be feared and/or avoided.

### An Argument for the Evolution of Human Morality

For the purposes of this paper, moral behaviour will be considered roughly equivalent to altruism, which is defined as the conferment of some benefit onto some other non-related conspecific organism at some cost to the actor (Trivers, 1971). On this definition, we would consider a man rescuing a drowning swimmer to be an altruistic act only if the swimmer was not his child, or otherwise related to him by blood. Examples of this type of behaviour are abundant in modern society, and they need not be acts of heroism. Everyday, mundane actions such as one person helping his friend move to a new apartment, or lending his friend money, or even just providing a simple favour such as a ride to the airport are all acts of altruism that are considered the “right” course of action when a friend is in need. All of these examples have one thing in common: the cost of the altruistic action to the actor is less than the benefit to the friend. So why would anyone ever bother going out of their way to help someone else when it costs them valuable resources such as energy and time, that could have otherwise been used perusing their own interests? Trivers (1971) suggests three conditions under which this type of reciprocal altruism would be expected to flourish: First, an abundance of opportunity for altruistic acts over

the lifetime of the individual must exist. Secondly, small groups of individuals must interact repeatedly, and finally, pairs of altruists must be exposed to roughly symmetrical altruistic situations in which they are able to confer similar advantages on each other for comparable costs. Trivers (1985, as cited in Krebs, 1998) concludes that human life coincides with these criteria perfectly:

During the Pleistocene, and probably before, a hominid species would have met the preconditions for the evolution of reciprocal altruism; for example, long life span, low dispersal rate, life in small, mutually dependant and stable social groups, and a long period of parental care leading to extensive contacts with close relatives over many years. (p.386)

The idea that such behaviours in humans could not have been created by the process of evolution by natural selection is usually based on the false idea that truly moral or altruistic actions would not yield the acting organism greater reproductive advantages (Krebs, 1998). However, when the alternative is considered, it too seems to contradict common sense. It is not the case that the most overtly selfish people yield greater benefits than those who chose other “moral” strategies. For example if you were to help a friend move to a new apartment, but then received no help in return when a similar need arose for you, your friend would be considered to be “wrong” in not repaying the implicit debt that he owes. In effect, he would be “cheating” in the sense that he failed to reciprocate, and would probably not receive further help from you in the future. The key point here is that the benefits that you would have conferred onto your friend would have been necessary for his survival in the evolutionary environment. Thus behaving in ways that are perceived as cheating would be a detriment to his survival and reproductive goals.

So, if cooperation was needed to survive and consequently, to reproduce in the evolutionary environment, then cooperators have must have had a selective advantage over non-cooperators. There is significant evidence pointing to the fact that cooperation was indeed a

necessary condition for survival for our ancestors. For example, hunting large game animals on the flatlands required a group effort, while only one hunter was responsible for the eventual kill. Hunters that shared their extra spoils with others were guaranteed food from other hunters who made kills in the weeks and months to follow. However, those hunters who endeavored to “cheat” and keep all the meat for themselves, not only wasted much of it to rotting, but also assured that they would be hungry in the weeks to come as other hunters would not share their meats with non-cooperators (Buss, 2004). In this way, natural selection has shaped our altruistic actions, at least in terms of sharing, over tens of thousands of years.

So, evolution may have shaped our tendencies to share in reciprocally altruistic ways when interacting with a small number of familiar people, but how, if at all, does it explain the larger moral processes in society such as justice, and our collective notions of right and wrong? To answer this question, we must first summarize the differential strategies that can be employed in moral dilemmas to maximize personal gain. Krebs (1998) lists the three key strategies that are used differentially according to circumstance: being moral when it pays off, cheating when it can be gotten away with, and punishing other cheaters when it is advantageous. These three strategies play out within the paradoxical theatre of reality in which greater cooperation yields the best average result for all parties involved, yet greater cooperation also entails greater vulnerability to cheating, whereby cheating becomes an increasingly tempting option over cooperation. This situation is perfectly analogous to the prisoner’s dilemma in which the best overall outcome for both prisoners occurs when each prisoner cooperates with the other, even though the optimal outcome for each individual prisoner occurs when one prisoner effectively takes advantage of the other by defecting, when the other cooperates (Trivers, 1971). This asymmetrical state of costs and benefits makes it beneficial for both prisoners to cooperate with

each other, especially if they often find themselves in such a dilemma, thereby allowing their mutual benefits to accumulate over time. Analogously, an example of one such mutual benefit in modern society is the division of labour, whereby more work can be accomplished overall, if individuals specialize in their roles, and can be trusted to trade their goods and services fairly (Ridley, 1997). As Adam Smith (1759, as cited in Pinker, 2002, p. 161) noted:

It is not from the benevolence of the butcher, the brewer or the baker that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love.

Boyd and Richardson (1992), claim that this type of large scale, or societal cooperation can be accounted for by Krebs' third strategy for maximization of personal gain: punishment of non-cooperators. They argue that "If the costs of being punished are large enough, *moralistic* strategies which cooperate, punish non-cooperators, and punish those who do not punish non-cooperators can be evolutionarily stable" (p.171). Furthermore, they claim that this strategy of punishment can be effective in producing mutually beneficial behaviours even in sizable groups.

But what about the popularly believed (and intuitive) notion that people help and support friends and family members only because it feels right to do so, not because they expect some future reciprocation? Certainly our emotions must figure into the mix; are the ideas of reciprocal altruism and selfish justice just too cold and calculated to be true? Dennett (2006) uses the example of the co-evolution between humans and fruit to demonstrate how the cold processes of "tit-for-tat" reciprocation and the warm fuzzy feelings of interaction with friends and family may indeed share the same underlying process. He explains that although there is nothing inherently "sweet" about the sugar molecules found within fruit, there certainly was a benefit for those who preferred and ate them in the past. Over tens of thousands of years, natural selection would favour those who were motivated to eat such sources of energy, and thereby favour those who

had an innate desire for such fruits. In effect, natural selection has hardwired the motivation to eat such fruits into the brains of our ancestors, and consequently, into us. So although there is an ultimate (i.e. why?) explanation for why we are motivated to eat fruit, it is also perfectly legitimate to explain it in terms of proximate (i.e. how?) causes such as motivations and emotions. As Dennett notes:

People generally say that we like some things because they are sweet, but this really puts it backward: it is more accurate to say that some things are sweet to us because we like them! (p. 59)

This reasoning, when applied to our moral motivations, serves to unify our subjective emotions and the cold logic of our objective evolutionarily stable strategies of altruism and justice. Trivers (1971) extends this logic one step further by arguing that emotions such as, friendship, gratitude, sympathy, guilt, trust, and suspicion, in conjunction with specific cheater detection, and cost benefit ratio mechanisms were selected for in order to uphold such systems of cooperation and justice amongst groups of humans. Trivers (1985, as cited in Krebs, 1998) concludes:

A sense of justice has evolved in human beings as the standard against which to measure the behaviour of other people, so as to guard against cheating in reciprocal relationships (p. 388).

In short, it may be the case that you cooperate with your friends and abide by the law because you “feel” that these are the “right” things to do, but it is also the case that natural selection has arranged it as such: on average, your emotional leanings towards moral conduct coincide quite nicely with the most efficient way to remain alive and maximize your personal gains.

Taken together, the above evidence suggests that contrary to popular belief, human morality is perfectly explicable by the processes of evolution by natural selection. The question of why this notion runs counter to popular belief is the focus of the following section.



### The Fear of Moral Imperfection

In light of all of this evidence (and much, much more), why do some people insist that the origins of morality lie outside the explanatory bounds of evolutionary theories? Pinker (2002) suggests that people fear that because the process of evolution is “red in tooth and claw”, it somehow discredits our hopes for consistent moral behaviour, and ultimately promotes violence and selfishness. Pinker has aptly named this the fear of “imperfectibility”: If it is the case that people are immoral by nature, would it not follow that attempts to improve our moral condition are frivolous? Are we doomed to our violent and selfish ways? A classic instantiation of this fear is Stephen J. Gould’s rhetorical question: “Why do we want to fob off responsibility for our violence and sexism upon our genes?” (as cited in Pinker, 2002, p.133). This fear was also painfully evident in 2001 when American congressman Tom DeLay literally blamed the Columbine high school massacre on a naturalistic understanding of human morality. He claimed that violence would certainly persist in America as long as “our school systems teach children that they are nothing but glorified apes, evolutionized out of some primordial soup of mud” (Falk 2001, as cited in Pinker, 2002, p.129).

It seems counterintuitive that any theory about origins could actually cause such violence and hatred. Indeed, the fear of imperfectibility is held aloft by an almost imperceptible logical fallacy that is floating just beneath the surface in all discussions of evolution. The is-ought problem was first identified by David Hume (1739). He correctly points out that prescriptive statements do not logically follow from descriptive ones. That is, regardless of how many descriptive statements one makes about how the world “is”, no prescriptive statement about how the world “ought” to be can follow directly. Note how the theory of evolution by natural selection as described and argued for above seeks to make descriptive statements about how the

world “is”. It does not seek to (nor could it ever for that matter) make prescriptive statements about how we ought to treat each other or how we ought to make laws that govern social behaviour.

This is the misunderstanding that makes evolutionary theory dangerous, especially when applied to human evolution and morality. Misinterpretations along these lines have led to atrocities such as social Darwinism, as well as eugenics programs such as that employed by the Nazis (Pinker, 2002). Without delving into the details of these atrocities, which fallaciously used evolutionary theory as a guise for sinister political agendas, it is sufficient to say that it simply does not follow logically that because certain forms of social cheating were useful or even required at some period in our evolutionary past, that robbing a bank, or stealing a car, or killing a neighbor is the “right” thing to do. To come to a sense of “right” and “wrong”, we must combine the objective facts that we know of the world with subjective value judgments about what we want the world to be like (Pinker, 2002). Theories concerning the evolution of human morality purport to alter our facts, not our values.

### Moral Progress

So perhaps we need not fear the total collapse of moral values, but what about the status of moral progress? How can we ever hope to improve our moral condition if we are all stuck with a static set of evolutionarily derived moral strategies? Robert Wright (2000) argues that increasing cooperation over time may in fact be an inevitable consequence of these static moral strategies, because they lead us to participate in non-zero-sum games (i.e. games where each player is better off for participating, as opposed to zero-sum games in which there is a winner and a loser). He claims that three aspects of our evolved human nature are responsible for putting

us on a “moral escalator” that has been responsible for increasing human cooperative efforts over time. Firstly, our ability to figure out how the world works has allowed us to accumulate technologies, both primitive and modern, that serve to provide greater opportunities for gains in trade. Secondly, the development of human language has aided in the trading of these technologies by allowing a clearer understanding of the terms of trade, as well as the threats of punishments for cheating. Finally, our moral emotions, such as friendship, gratitude, sympathy, guilt, trust, and suspicion as discussed above, allow human cooperators to effectively mediate relationships with current cooperators, as well as seek out relationships with new cooperators. Thus, Wright argues, that because you cannot both kill someone and trade with them, the group of people that we consider to be worthy of trust and respect can actually increase with the number of people that we trade and cooperate with. The system is comparable to an escalator because of its tendency to act like a positive feed-back mechanism: The more groups cooperate with each other, the more technology becomes available, and these new technologies serve to further increase cooperation between such groups. So what does this have to do with global cooperation in terms of avoiding conflicts and preventing wars? Wright (2000) points out that human societies are subject to the same laws of cooperation that individual organisms are. So long as cooperating entities can eliminate the problems of punishing cheaters, it will always be the case that these entities will be better off when they cooperate in the pursuit of common goals by dividing labour where possible and thereby accumulating gains in trade. As Wright so gently puts it: “Among the many reasons I don’t think we should bomb the Japanese is that they built my minivan” (2000, as cited in Pinker, 2002, p. 320). Thus, because our fixed moral sense, endowed to us through the process of evolution, implores us to take part in cooperative non-zero-sum games, genuine moral progress is indeed possible, even at a global level.

### Conclusion

I have argued that not only is human evolution perfectly explicable as a product of evolution by natural selection, but also that genuine moral progress can proceed on a global scale as a result of this fixed moral sense contributing to our motivation to participate in non-zero-sum games. Based on the evidence for these claims, it seems that not only can we trust in our biologically based sense of morality to guide us through everyday interactions in the present, but we can also look forward to genuine moral progress as a society, and as a planet, well into the future as a result of its design. If these claims are true, then it is also true that the fear of moral imperfection attributed to a biological understanding of human nature and morality is simply an overly pessimistic view, perpetuated by short sightedness and an underlying fear of the future. In light of these facts, it is simply not the case that a biological understanding of human morality or human nature ought to be feared and/or avoided.

## References

- Boyd, R. & Richardson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13, 171-95.
- Buss, D.M. (2004). *Evolutionary psychology: The new science of the mind*. Boston: Pearson Education Inc.
- Dennett, D.C. (2006). *Breaking the spell: Religion as a natural phenomenon*. New York: Viking.
- Falk, D. (2001, October 21). Design or Chance? *Boston Globe Magazine*, 14-23.
- Gould, S.J. (1976). Criminal man revived. *Natural History*, 85, 16-18.
- Hume, D. (1739/2000). *A treatise of human nature*. New York: Oxford University Press
- Kohlberg, L. (1984). *Essays on moral development: Vol. 2. The Psychology of moral development*. San Francisco: Harper & Row.
- Krebs, D.L. (1998). The evolution of moral behaviours. In Crawford, C., & Krebs, D. L. (Eds.), *Handbook of evolutionary psychology: Ideas, issues, and applications*. Mahwah, N.J.: Lawrence Erlbaum.
- Pinker, S. (2002). *The blank slate: the modern denial of human nature*. London: Penguin.
- Ridley, M. (1997). *The origins of virtue: Human instincts and the evolution of cooperation*. London: Penguin.
- Smith, A. (1759/2002). *The theory of moral sentiments*. Cambridge: Cambridge University Press.
- Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.
- Trivers, R.L. (1985). *Social Evolution*. Menlo Park, CA: Benjamin/Cummings.
- Wright, R. (2000). *Nonzero: The logic of human destiny*. New York: Pantheon.