

A Review of Matt Ridley's
THE ORIGINS OF VIRTUE:
HUMAN INSTINCTS AND THE EVOLUTION OF COOPERATION*
by Jack Hirshleifer

Working Paper Number 771
Department of Economics
University of California, Los Angeles
Bunche 2263
Los Angeles, CA 90095-1477
August 1997

*Prepared for REASON Magazine.

A Review of Matt Ridley's

THE ORIGINS OF VIRTUE:

HUMAN INSTINCTS AND THE EVOLUTION OF COOPERATION*

Abstract

The paradox addressed by Ridley is how to reconcile the extent of cooperation among humans with the "selfish gene" postulated by sociobiology. Apart from mutual aid that is merely incidental to pursuit of self-interest, or that can be attributed to kinship, Ridley attributes observed cooperation on the human level to a single force: reciprocity. He argues that reciprocity is not just sound business calculation, but is supported by evolved human instincts. Granted these instinctive motivations, government is supposedly unnecessary. But reciprocity alone cannot bear the weight that Ridley places on it; the problem remains of controlling thieves and cheaters. Absent external enforcement through government, a different kind of virtue is required: willingness to engage in disinterested moral aggression against violators. Such an instinct could have arisen by cultural evolution or by group selection.

Sociobiology, the research program that inquires into the extent and limits of evolutionary influences upon behavior, has been fortunate in its expositors. Apart from Edward O. Wilson's magisterial foundation text *SOCIOBIOLOGY* (1975), a selection of other important contributions would include Michael Ghiselin's *THE ECONOMY OF NATURE AND THE EVOLUTION OF SEX*, Richard Dawkins' *THE SELFISH GENE* and his *THE EXTENDED PHENOTYPE*, David Barash's *THE WHISPERINGS WITHIN*, and Matt Ridley's own earlier *THE RED QUEEN*. The growing intellectual influence of sociobiology has redeemed the "nature" portion of the nature-versus-nurture controversy from the anti-biologism of once-reigning social science figures such as B. F. Skinner, Franz Boas, and Margaret Mead.

I hasten to add that speaking of nature versus nurture -- setting the genetic endowment in opposition to cultural influence -- is seriously misleading. Almost all the phenomena needing explanation involve both elements, interwoven in subtle and fascinating ways. Our ability to pick up languages in the early years of life is a genetically ingrained capacity, but whether a child learns English or Chinese or Swahili depends upon the cultural environment. Or to take an issue closer to the theme of the book under review, people have an innate need to affiliate with others, to distinguish between "us" and "them". But whether the in-group and out-group are defined in terms of geography or skin color or religion or language is mainly a matter of cultural influence.

In *THE ORIGINS OF VIRTUE*, Ridley deals with the natural selection of human instincts promoting the evolution of cooperation. (Though I do not have space for a comparison here, the topic has also been addressed in two other recent excellent works: Robert Wright's *THE MORAL ANIMAL* and Frans de Waal's *GOOD NATURED: THE ORIGINS OF RIGHT AND WRONG IN HUMANS AND OTHER*

ANIMALS.) Cooperation poses an intellectual problem, since the driving principle of sociobiology is "the selfish gene" -- that each and every gene has been selected over evolutionary history to be good at one thing above all else: promoting its own reproductive survival and multiplication. But genes do not float nakedly around the biosphere. They have linked up into chromosomes, and chromosomes come together in the genomes of living organisms. Thus each gene must somehow meld its own reproductive survival with that of neighbors and partners. "Selfish genes sometimes use selfless individuals to achieve their ends" (p. 20). And this tension between cooperation and conflict in what Ridley calls the society of genes is reflected by the mix of selfish and unselfish drives and behaviors on the level of organisms:

Our minds have been built by selfish genes, but they have been built to be social, trustworthy and cooperative. That is the paradox this book has tried to explain. (p. 249)

Still, without espousing the extreme culture-is-all position of Boas, a critic might query the need for genetic explanations of morality or virtue. Even some sociobiologists, among them Richard Dawkins and George C. Williams, place morality in opposition to our inherited natures, achievable only if human culture allows us to overcome our selfish genes.

Ultimately, perhaps, genes for love or for mutual aid may actually be identified along this or that chromosome. Absent such direct proof, it remains plausible to conjecture an evolutionary origin if (for example) a behavior approaches universality in the human species, if it is easy to learn and hard to eradicate, and if parallels or precursors can be identified among our cousins the apes. On these criteria, the case for a

genetic source of morality is somewhat mixed. True, such virtues as courage and generosity and marital fidelity are honored in just about every human society. Honored yes, but practice falls far short. Easy to learn as precepts yes, but easy to forget when it comes to deeds.

On the other hand, the very fact that humans find it advisable at least to counterfeit virtue is instructive. Pretence would be pointless if the real thing did not exist. As for continuities with other species, in his GOOD NATURED the primatologist Frans de Waal declares: "Are animals moral? Let us simply conclude that they occupy a number of floors of the tower of morality." Overall, therefore, at least for the purpose of this review I will provisionally accept Ridley's underlying presupposition: that many of our pro-social behaviors do reflect ingrained universals based upon the evolution of the human species.

Some forms of cooperation, being merely incidental to self-interested actions, do not pose any special explanatory problem. If predator efficiency is lower when prey flock together, no special virtue attaches to a prey animal who joins a flock and thereby helps all the others as well. (Sociobiologists call this the "selfish herd" model.) Or consider conventions. If everyone else is driving on the right, even in the absence of policing it would not make sense to deviate and drive on the left.

Almost as easy to explain, at least for sociobiologists, is "nepotism" -- cooperating with kin. Helping one's relatives is, from the point of view of the genes, in large part self-help. Working out the details as to why some species display parental love and care while others do not, and why even in a given species there may be huge variations in response to environmental contexts and cues, is a main subject of sociobiological

research.

Ridley devotes little space to these easy cases, turning quickly to the harder, more important instances: helpful acts that are not incidental by-products but require real self-sacrifice, and where kinship is not a factor. His explanation can be summarized in one word: reciprocity. Sacrifices on behalf of non-kin are undertaken only for compensation. (I'll scratch your back if you'll scratch mine.) Willingness to reciprocate is often just good business sense, but Ridley finds its source in something deeper:

/R/eciprocity may be an inevitable part of our natures: an instinct.

We do not need to reason our way to the conclusion that 'one good turn deserves another'. (pp. 65-66)

In Chapter 2, Ridley convincingly argues that specialization and the division of labor among humans -- among animals, only the social insects can compare -- are not evolutionarily recent consequences of urban civilization or the Industrial Revolution. Rather, they have been around long enough to be ingrained into the human psyche. A 5000-year old Neolithic man has been found buried with an astonishing variety of clothing, weapons, and implements of metal, bone, and textile far beyond what a single person could himself have fabricated or appropriated. His panoply was almost surely accumulated by reciprocal exchanges within his band, possibly even among different bands.

While in sympathy with Ridley's enthusiasm, I see two difficulties. First, where is the virtue here? Most of us identify "true" virtue with unreciprocated generosity, for example the soldier who throws himself on a live grenade to save his comrades. Willingness to be helpful for payment is at best a pallid sort of morality. And second, can a social order based

solely upon reciprocal exchanges really be viable, or must it be founded upon something more truly virtuous?

Taking up the latter question first, social systems based upon reciprocity are subject to subversion. Exchange presumes a pre-existent structure of property rights, an accepted mine and thine. But there always have been and always will be robbers ready to seize others' possessions. In a slightly different category are cheaters who fail to pay for benefits received. In modern times we call upon the law to keep such malefactors under control: to define and defend property rights and to enforce reciprocal agreements. For Ridley, however, trade precedes law (p. 202). The problem is, absent law, how are rogues and scoundrels to be controlled?

To fend off robbers, is each person to protect his possessions by force? Granted, self-defense is not immoral, but among other objections it has the undesirable feature of making each man judge in his own cause. Ridley does not systematically address the problem, but to my mind he is on dangerous ground in approving the "harbor gangs" who have established their own de facto property rights in the Maine lobster fisheries. ("Although it is illegal to cut somebody's traps free from their buoys, it happens regularly to any intruder." -- p. 229.)

Ridley does deal at length with the other main type of subversion, failure to reciprocate. Chapters 3 and 4 are devoted to his claim that on this score self-enforcement will suffice. Having encountered a non-reciprocator, a person could simply decline further business interactions. This is the essence of what has become known as the TIT FOR TAT strategy. TIT FOR TAT never earns short-run gains by exploiting those willing to cooperate, but its refusal to do further business with defectors means it

cannot be exploited too severely either. Supposedly, TIT FOR TAT not only controls defection but does so better than any other strategy, in fact will drive all other strategies to extinction in the evolutionary long run.

This claim, though often repeated in the sociobiological literature -- like people in every field, sociobiologists mainly talk to and echo what they hear from one another -- is incorrect. Game theorists know that sometimes TIT FOR TAT does OK, sometimes not; it all depends upon the specifics of the situation.

Ridley is not unaware of the difficulty, and several of the most interesting passages in the book deal with it. TIT FOR TAT would be more successful, evidently, if people could identify potential defectors in advance -- and there is some evidence that they can! Also, emotions like anger may have evolved to discourage cheaters, by eliciting disproportionate ("irrational") punishment reactions on the part of victims.

Still, reciprocity alone really cannot support the weight that Ridley would have it bear, cannot of itself explain the extent of cooperation among non-kin. A system of exchange based upon property rights must rest upon more than self-defense and TIT FOR TAT responses. And in particular, disinterested third parties have to be willing to engage in what has been called "moralistic aggression" to defend victims and punish defectors. If so, reciprocity is not the origin of virtue. Rather, true morality -- pro-social propensities motivated by principle or compassion rather than by expected compensation -- must be there already if a system of trade and exchange is to be viable.

Granted that true virtues, extending upon occasion to heroic acts of self-sacrifice, do exist and cannot stem from mere reciprocity, how can they

have evolved? In *THE MORAL ANIMAL*, Robert Wright had the delightful idea of using the life of Charles Darwin himself to illustrate the role of ingrained human instincts in helping resolve ethical quandaries and moral choices. In his concluding pages Wright asks whether a truly civilized, generous, idealistic, honorable individual like Darwin can really be explained in terms of his selfish genes. The final conclusion: in the light of the utter ruthlessness of evolutionary logic, Charles Darwin's morality -- and indeed ours, such as it is -- remains "nearly miraculous".

There are ways out of the puzzle. As suggested above, morality might be a human cultural development. By rising above our ingrained animal natures, by following the counsels of our sacred texts and great sages, we may have partially succeeded in a "revolt against the selfish genes". An alternative, more mundane possibility was suggested by Darwin himself:

When two tribes of primeval man, living in the same country, came into competition, if (other circumstances being equal) the one tribe included a great number of courageous, sympathetic and faithful members, . . . , this tribe would succeed better and conquer the other.

(Darwin, *DESCENT OF MAN*)

Darwin is suggesting that the habit of cooperation, even where disadvantageous to the individual, might be evolutionarily viable owing to differential group survival. Darwin's idea here has come to be known as "group selection", a concept currently scorned by majority opinion among sociobiologists (and by Ridley, Wright, and de Waal as well). There is however an important dissenting minority; a forthcoming book by David S. Wilson and Elliott Sober will be making a strong case for group selection. The issues involved are too technical for analysis here, but to my mind the

evidence for the power of group selection -- particularly in shaping the behavior of intensely group-competing species like ants and humans -- seems overwhelming.

Going back to TIT FOR TAT, the intellectual enterprise of looking for a single "best" strategy is seriously misguided. What is best at any moment of time depends upon the details of the situation and upon the other strategies in play. Accordingly, what we observe is not the universal adoption of any single strategy but rather an ever-changing co-evolving mixture. In fact, as shown in some remarkable simulations by Bjorn Lomborg, viable patterns of cooperation practically always involve combinations of strategies. Producers, cheaters, and what we might call guardians (who keep cheaters under control) are typically all present simultaneously in evolving populations. This is closer to the way cooperation, halting and imperfect yet nonetheless real, is achieved in our fallen world.

Readers of this magazine will find Ridley's strong pro-market stance, set forth mainly in his final three chapters, appealing. But most pro-market conservatives do admit, despite the evident dangers, a substantial role for government -- to maintain law and order, protect property rights, and enforce contracts. As opposed to this "classical liberal" orientation stemming from John Locke, Adam Smith, and our constitutional founding fathers, Ridley leans in the anarchist direction. Individuals, motivated solely by the expectation of reciprocity, can supposedly do it all themselves -- establish property rights, produce, trade, punish cheaters -- if only government will get out of the way.

Yet systems of dominance and subordination, which are the essence of government, are also in our evolutionary heritage, as explored notably in de

Waal's GOOD NATURED and Roger Masters' THE NATURE OF POLITICS. The possible effectiveness of government for good probably rests upon certain other human virtues -- among them loyalty and noblesse oblige -- that call for further examination. But that is the topic for another book.

[about 2350 words]