"Must there be human genes specific to prosocial behavior?"

Michael D. Cohen, Robert Axelrod, and Rick Riolo
University of Michigan

{mdc, axe, rlriolo}@umich.edu

June 27, 2002

Joseph Henrich's essay on "Cultural Group Selection…" makes many insightful observations about evolutionary models of cooperative behavior. At the same time, we think there are several important points in his presentation that are questionable, and we don't believe he has made the case for one for his key conclusions: that cultural processes have established – in humans alone – genetic traits for forms of behavior such as ingroup altruism and punishment of deviants. We first set out our reconstruction of his argument; then we offer commentary on each of its components.

Henrich offers a series of five assertions that, in our reading of his essay, provide an outline of this central argument.

1. Two facts worth explaining are that (a) on average, human beings exhibit much higher levels of cooperation than other species, except perhaps prosocial insects; and (b) particular cooperative behavior patterns vary widely among human subgroups.
2. Models seeking to account for these facts strictly from genetic processes are inadequate and unlikely to be sufficiently improved.
3. Therefore, models are needed that bring into focus the interaction between cultural processes and genetic ones.
4. In such models, one can expect group selection dynamics to be powerful and to lead to group selection at the cultural level, as well as the genetic level.
5. In this setting, some groups will have prosocial practices, and in such groups there will be favorable conditions for establishing prosocial genes for such behaviors as "ingroup altruism and punishment [of deviators]" (manuscript p. 40).

We respond to these five assertions in turn.

With regard to the first assertion, we agree that humans exhibit unusually high levels of cooperation (as well as conflict). We also agree that cooperative behavior patterns vary widely among human subgroups.

Henrich's second assertion is based on a critique of mechanisms that have been proposed to account for prosociality. He groups them together as "genetic" although we see no reason why reciprocity, for example, must be founded on specifically prosocial genes. We think Henrich is quite correct to point out that a special and fairly strong assumption is crucial to the performance of the "green beard" model, namely that some recognizable, distinctive feature of the organism is almost perfectly correlated with prosocial behavior. This does limit the applicability of that model to cases corresponding to this assumption.

In later subsections he points out limitations in the valid scope of other models. But he tends to imply that any model that fails to explain everything explains nothing. He discards each successive model because he finds something it cannot explain or because it incorporates a constraint he deems inappropriate.

This unsound rhetoric leads to trouble. In the end, he is not correct in concluding from his review of the greenbeard and other models that *no* model based on weaker, more plausible assumptions can be successful across a useful range of cases. Despite his skepticism, more plausible mechanisms of this kind have been shown theoretically capable of generating large-scale prosociality in species far simpler than humans. Our recent work on a tag-based mechanism for the "donation game" is a case in point (Riolo, Cohen and Axelrod, 2001; see also the commentary by Sigmund and Nowak, 2001). We show that randomly assigned visible labels are sufficient scaffolding for a population of adapting agents to acquire strategies of making costly donations to others they will probably never meet again. This is precisely the kind of one-shot, anonymous situation for prosocial action that Henrich claims cannot be explained with simple models that do not rely on culture-gene coevolutionary processes. In our model, such cooperating populations do experience periodic breakdowns as egoists prosper by receiving without giving, but cooperative behavior quickly reemerges around some other detectable tag. Because tag-based mechanisms can be found in both social and biological systems, the requisite formation of cooperative "groups" can be endogenous to simple social or biological processes (Cohen, Riolo and Axelrod, 2001). Models such as these need not rely on biological group selection or on culture-gene coevolutionary processes that result in genes specific to prosociality.

While we don't think Henrich has established that models of direct and indirect reciprocity will never explain a significant portion of cooperation, we do think there is much good work to be done along the lines referred to in his third assertion, namely studies of the interaction of culture and genes.

We also agree with much of his fourth assertion, namely that one reason for stronger "group selection" is that cultural factors strongly bias the interactions among agents in a population, altering what Axelrod and Cohen (1999) have called the "interaction topology" of the system. This can allow between-group differences to be amplified relative to within-group differences. We disagree, however, with Henrich's minimization of the differences between how genes and culture are transmitted. These differences are especially important after the vast changes he notes in both what counts as "fitness" and what is being replicated.

The fifth and final assertion of Henrich's argument is that a world with widespread "cultural group selection" will have some strongly cooperative subgroups that provide fertile ground for genetically establishing traits that are specific to prosocial behaviors. He mentions ingroup altruism and punishment of deviants, and suggests that there may be others, such as "abilities to rank potential models according to their payoffs, and to preferentially imitate highly ranked models" (manuscript p. 30).

In our view, Henrich gives insufficient weight to more general properties of human beings, such as language and foresight, that have enormous survival value for many other reasons, but that are also mobilized in our cooperative institutions and practices. When he does touch on "social learning" and "quality imitation" he bypasses important literature. As early as 1922 John Dewey developed an evolution-based view of docility (teachability) as the cornerstone of human cultural transmission and social institutions. Indeed, Dewey's analysis is more nuanced, taking into account that docile individuals can also be exploited by their "teachers". Herbert Simon (1990) makes similar arguments regarding the value of docility and goes on to work out the mathematical conditions for the advantages of social learning to outweigh concomitant exploitability. These approaches are among many that account for human patterns without invoking the evolution of prosocial genes, as argued in Henrich's fifth assertion.

As Henrich points out, humans can sometimes arrange cooperation with others who are not members of their own group. To account for observed patterns of cooperation, he invokes the existence in humans of specifically prosocial genes that can support behavior such as ingroup altruism and punishment of deviators. While such specialized prosocial genes are possible, we are not convinced they are necessary to account for observed behavior. Instead, we think a more promising approach is to examine the nature of cooperative institutions and practices, and focus directly on the logic of social interaction among agents with general-purpose capabilities such as language and foresight.

WORKS CITED

Axelrod, R., Cohen, M. D., 1999. Harnessing Complexity: Organizational Implications of a Scientific Frontier. The Free Press, New York.

Cohen, M. D., Riolo, R., Axelrod, R., 2001. The Role of Social Structure in the Maintenance of Cooperative Regimes. Rationality and Society 13, 5-32.

Dewey, J., 1988. Human Nature and Conduct (1922). In J. A. Boydston, (Ed.), The Middle Works of John Dewey, Vol. 14. Southern Illinois University Press, Carbondale and Edwardsville, Illinois.

Riolo, R., Cohen, M. D., Axelrod, R., 2001. Evolution of Cooperation without Reciprocity. Nature 414, 441-443.

Sigmund, K., Nowak, M., 2001. Evolution: Tides of Tolerance. Nature 414, 403-405.

Simon H. A., 1990. A Mechanism for Social Selection and Successful Altruism. Science, New Series, Vol 250, Issue 4988, 1665-1668.