

The Social Structure of Cooperation and Punishment

Herbert Gintis and Ernst Fehr

April 26, 2011

Standard models of human cooperation in economics and biology assume purely self-regarding agents who used repeated interactions (reciprocal altruism) or public reputations (indirect reciprocity) to sustain mutual helping behaviors. While these mechanisms are important, there are many equally important forms of prosocial behavior that cannot be account for thereby (Gintis 2005, Fehr and Gintis 2007, Gintis 2009, Bowles and Gintis 2011). These include voting in elections, participating in collective actions, being kind to strangers, contributing to community public goods, and behaving morally in anonymous situations, or where the material penalties for immoral behavior are low.

Economic experiments strongly suggest that human prosociality is not limited to calculated selfishness (Fehr and Tyran 1996, Fehr et al. 1997, Fehr et al. 1998, Fehr and Gächter 1998, Fehr and Gächter 2000b, Fehr and Gächter 2000a, Fehr and Gächter 2002), but that the presence of free-riders is a key and ever-present threat to sustained cooperation. *Strong reciprocity*, a behavioral mechanism including both altruistic cooperation and costly punishment of free-riders (Gintis 2000) thus helps sustain cooperation over long periods. This work showed that humans have strong and consistent other-regarding preferences that could be enlisted in support of social cooperation. In fact, anthropologists have confirmed that strong reciprocity is indeed routinely harnessed in the support of cooperation in small-scale societies (Boehm 1984, 2000, Wiessner 2005, 2009, Henrich, Ensminger et al. 2010), as stressed in Joseph Henrich's commentary in this issue.

Guala characterizes the punishment side of strong reciprocity as “uncoordinated.” This is simply incorrect. Collective action is a real-life expression of strong reciprocity (Bowles and Gintis 2004, p. 17) and the predisposition to punish “transgressors” is often socially organized and sanctioned. Indeed, individuals are often deterred from carrying out self-initiated sanctions (Boyd et al. 2010). The experimental evidence for coordinated punishment was laid out in several experimental papers on strong reciprocity (e.g., Cinyabuguma, Page, and Putterman 2005).

Guala claims that costly punishment is rarely observed in the real world, and what punishment is observed is generally not very harsh (e.g., verbal harassment,

gossip, ostracism). These observations, even if true, in no way conflict with strong reciprocity models of social cooperation. First, if punishment is effective, it will be rarely carried out. Thus, the absence of frequent punishment is indication that the threat of punishment has a particularly strong effect. For instance, the average tax payer in the United States is never penalized for tax evasion, yet no one doubts the importance of prosecuting tax evasion. Similarly, most drivers receive only a few traffic citations in the course of their lives, but many drivers adjust their driving to avoid citations. Second, we stress that most humans are very averse to public criticism of even a verbal form of punishment, and we cite studies that show that verbal criticism alone often leads to conformity (Masclot et al. 2003). In addition, the human emotion of shame serves to amplify social criticism, thereby lessening the need for costly punishment (Gintis 2004, Bowles and Gintis 2005). Moreover Guala seriously understates the importance of diffuse, uncoordinated, costly punishment in promoting norm-adherence.

Guala claims that some punishment is zero cost. If so, this would add an interesting dimension to the strong reciprocity model, but it does not conflict with this model.

In sum, we agree with Guala that social structured punishment is important, but we assert that the predisposition to reward goodness and punish evil underlies the effectiveness of socially structured punishment. We also reaffirm the critical importance of diffuse, unstructured, cooperation and punishment in fostering social efficiency and a high quality of life.

REFERENCES

- Boehm, Christopher, *Blood Revenge: The Enactment and Management of Conflict in Montenegro and Other Tribal Societies* (Philadelphia, PA: University of Pennsylvania Press, 1984).
- , *Hierarchy in the Forest: The Evolution of Egalitarian Behavior* (Cambridge, MA: Harvard University Press, 2000).
- Bowles, Samuel and Herbert Gintis, “The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations,” *Theoretical Population Biology* 65 (2004):17–28.
- and —, “Prosocial Emotions,” in Lawrence E. Blume and Steven N. Durlauf (eds.) *The Economy As an Evolving Complex System III* (Santa Fe, NM: Santa Fe Institute, 2005) pp. 339–366.
- and —, *A Cooperative Species: Human Reciprocity and its Evolution* (Princeton: Princeton University Press, 2011).

- Boyd, Robert, Herbert Gintis, and Samuel Bowles, "Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare," *Science* 328 (30 April 2010):617–620.
- Cinyabuguma, Matthias, Talbot Page, and Louis Putterman, "Cooperation Under the Threat of Expulsion in a Public Goods Experiment," *Journal of Public Economics* 89 (2005):1421–1435.
- Fehr, Ernst and Herbert Gintis, "Human Motivation and Social Cooperation: Experimental and Analytical Foundations," *Annual Review of Sociology* 33 (August 2007):43–64.
- and Jean-Robert Tyran, "Institutions and Reciprocal Fairness," *Nordic Journal of Political Economy* (January 1996):1–18.
- and Simon Gächter, "How Effective Are Trust- and Reciprocity-Based Incentives?," in Louis Putterman and Avner Ben-Ner (eds.) *Economics, Values and Organizations* (New York: Cambridge University Press, 1998) pp. 337–363.
- and — , "Cooperation and Punishment," *American Economic Review* 90,4 (September 2000):980–994.
- and — , "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14,3 (Summer 2000):159–181.
- and — , "Altruistic Punishment in Humans," *Nature* 415 (10 January 2002):137–140.
- , — , and Georg Kirchsteiger, "Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica* 65,4 (July 1997):833–860.
- , — , Erich Kirchler, and Andreas Weichbold, "When Social Norms Overpower Competition—Gift Exchange in Labor Markets," *Journal of Labor Economics* 16,2 (April 1998):324–351.
- Gintis, Herbert, "Strong Reciprocity and Human Sociality," *Journal of Theoretical Biology* 206 (2000):169–179.
- , "The Genetic Side of Gene-Culture Coevolution: Internalization of Norms and Prosocial Emotions," *Journal of Economic Behavior and Organization* 53,1 (January 2004):57–67.
- , "Behavioral Game Theory and Contemporary Economic Theory," *Analyse & Kritik* 27,1 (2005):48–72.
- , *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton: Princeton University Press, 2009).
- Henrich, Joseph, Jean Ensminger, Richard McElreath, and Abigail Barr et al., "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment," *Science* 327 (19 March 2010):1480–1484.

- Masclot, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval, "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review* 93,1 (March 2003):366–380.
- Wiessner, Polly, "Norm Enforcement among the Ju/'hoansi Bushmen: A Case of Strong Reciprocity?," *Human Nature* 16,2 (June 2005):115–145.
- , "Experimental Games and Games of Life Among the Ju/'hoansi Bushman," *Current Anthropology* 50,1 (2009):133–138.

c:\papers\BBS Guala Commentary\GualaReply April 26, 2011