

*Behavioral and Brain Sciences* 28, 815-816, 2005

## **You can't give permission to be a bastard: Empathy and self-signaling as uncontrollable independent variables in bargaining games**

Commentary on Henrich, Joseph, Boyd, Robert, Bowles, Samuel, Camerer, Colin, Fehr, Ernst, Gintis, Herbert, McElreath, Richard, Alvard, Michael, Barr, Abigail, Ensminger, Jean, Henrich, Natalie Smith, Hill, Kim, Gil-White, Francisco, Gurven, Michael, Marlowe, Frank, W., Patton, John Q., and Tracer, David (2005) "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences* 28, 795-815.

by George Ainslie

151 Veterans Affairs Medical Center, Coatesville, PA 19320  
University of Cape Town, Rondebosch 7701, South Africa

**george.ainslie@va.gov**  
**www.picoeconomics.org**

*This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA, USA. The opinions expressed are not those of the Department of Veterans Affairs of the U.S. Government. It is considered a work of the U.S. government and as such is not subject to copyright within the United States.*

### **Abstract**

Canonical utility theory may have adopted its selfishness postulate because it lacked theoretical rationales for two major kinds of incentive: empathic utility and self-signaling. Empathy – using vicarious experiences to occasion your emotions – gives these experiences market value as a means of avoiding the staleness of self-generated emotion. Self-signaling is inevitable in anyone trying to overcome a perceived character flaw. Hyperbolic discounting of future reward supplies incentive mechanisms for both empathic utility and self-signalling. Neither can be effectively suppressed for an experimental game.

### **Text**

Henrich et.al.'s project has been an invaluable step forward in cross-cultural research, systematically collecting actual behavioral data in a design that makes data from diverse societies comparable. The result extends a finding in developed societies that is anomalous for utility theory as often interpreted: In ultimatum-type games people make offers that are greater than necessary to avoid rejection; and the risk of rejection is itself substantial even though the strictly monetary incentive is to accept any offer. The variability of responses among cultures is of great interest, of course, particularly the increase of altruism with degree of market integration until this altruism drops off sharply in western college students. However, it is also remarkable that this strenuous

undertaking should have been thought necessary to demonstrate the basic robustness of fellow feeling.

Economic man and his Darwinian kin have always represented somewhat of a victory of theory over common sense. Someone's refusal to share will generally strike an observer as callous or even hostile, and elicit a negative empathic response unless the person has "good reason." Utility theories that have felt compelled to view selfishness as rational in situations where cooperation cannot be rewarded have probably done so because they lacked concepts of either empathic utility or self-signaling. I will suggest that some form of these two concepts can let utility-based theory advance to the point of predicting what we have known all along.

Bargaining games try to create self-contained situations, so that subjects are motivated only by the contingencies of the game itself. They can only partially achieve this isolation. Rewards can be made large and the subjects can be tested in privacy, as in the research reported by Henrich et.al., but at least two major kinds of incentive remain uncontrolled. Perhaps not coincidentally, they come from the very sources that have been undeveloped theoretically, empathic utility and self-signaling.

**Empathic utility.** Although well-adapted individuals undoubtedly have to husband resources with which to propagate their genes for posterity, organisms do not think about this goal as such, and might not sympathize with it if they did. It is true that the reward process that evolution has shaped to give adaptiveness immediate salience includes the satisfaction of material needs, but it also involves emotions, which, whatever their adaptive function for the species, stand on their own as motivators of individuals. For instance, people not only give when they feel love and withhold when they feel hate, but they also find it rewarding to cultivate loves and hatreds. The occasions that support love and hate undoubtedly have some relationship with the circumstances where love and hate will bring materially adaptive outcomes, but only in a general way, in the big picture, where they select for the various degrees of preparedness our species has to generate these emotions. For the individual these emotions are consumption goods in their own right, to be sought and sometimes paid for in social organizations, taverns, movies, rehearsed memories, and fantasies – independently of whether they are otherwise profitable.

I have argued elsewhere that, because of our hardwired hyperbolic overvaluation of imminent experiences, emotions generated *ad lib* become relatively unrewarding. That is, our impatience for satisfaction makes us anticipate whatever is predictable, so that self-generated emotions pale into daydreams, creating the incentive to make vicarious experiences the most important occasions for our emotions (Ainslie 2001, pp. 161–86; in press, sects. 10 and 11.2). In addition, preliminary neurophysiological evidence suggests that vicarious experience may be generated readily at the most basic level: The brain motor area controlling a particular part of the body becomes active when a person sees someone else moving that part (Iacoboni et al. 1999). It looks like we are both innately prepared to model the experiences of others in ourselves and motivated to occasion our emotions with this model – unless, perhaps, as Henrich et.al. note, we are autistic. Thus,

the proposer in an ultimatum game will be aware of offering the other player not only money but also an emotional occasion, a choice that creates an emotional occasion for the proposer herself. She will feel generous or unremarkable or stingy, feelings that have values in themselves.

For the responder, if the occasion offered seems insulting, angry rejection may promise more reward than the prospect of having the money. It has always been clear that not all empathy is positive –for example, a person’s motive for retribution usually goes beyond the practical need for deterrence and involves an emotional appreciation of the target individual’s discomfiture (Leach et al. 2003; see, in the extreme case, Davies 1981, pp. 78-82). A possible mechanism by which empathy for pain can be satisfying is even more speculative than the one for vicarious reward itself (Ainslie 2001, pp. 183–86), but the idea that occasions for emotion can compete with money and other goods in the marketplace of utility should not be controversial.

**Self-signaling.** Beyond the imagined impact of their moves on other players, human subjects will also be concerned about what their choices tell them about their own characters. I have argued that hyperbolic discount curves for valuing future events innately dispose people to prefer poorer, earlier options to better, later ones *temporarily*, when the poorer options are imminently available (Ainslie 2001; in press). This disposition leads us to adopt devices to forestall these temporary preferences, the most powerful of which is the perception of current choices as test cases predicting entire bundles of future choices. This perception (I argue) is recognizable as willpower, which increases the force of the better options but makes our expectation of getting the bundle of better options vulnerable to any individual lapse. Proposers in ultimatum games might well see their choices as tests of whether they have overcome selfishness.<sup>1</sup> Responders might be concerned about overcoming either meekness or resentment. In such cases the subjects would have an incentive to avoid seeing themselves set a bad precedent, quite apart from the incentives created by the game. Of course, they might count their choices in a one-time experiment as exceptions to their resolutions; but if they had been giving their impulse a wide berth they might regard considerateness, say, to be a character trait, and count the making of any inconsiderate offer as symptomatic of not really having the trait. When I have run multi-person repeated prisoner’s dilemmas with an explicit instruction to try only to maximize individual winnings, subjects have cooperated when it was clearly not in their financial interest, and subsequently told me, “I’m just the kind of person who does that.” Bodner and Prelec discuss the wide-berth case, presumably refined thus from individual self-control efforts, in a paper on “self-signaling” (Bodner & Prelec 2001). So, arguably, does Max Weber in his explanation of why Calvinists’ belief in predestination seemed to increase their self-control (Weber 1904/1958; see Ainslie, 2001, pp. 135-136). If you are at pains to overcome a basic human urge in your life, you probably will not let an experimenter give you permission to indulge it, even “just this once.”

In Henrich et al.’s rich data we get glimpses of how cultural pressures shape people’s occasions for emotion and the “kind of person” they try to be. Orma subjects’ anonymous matching of progressive *harambee* contributions seems especially fine-tuned.

What we do not see, and will never see, is choice based entirely on the ostensible contingencies of reward in a bargaining game.

## Note

1. Determined canonical theorists might see such choices as a test of whether they had overcome “irrational” empathic urges – hence the reported epidemic of selfishness among economists (Frank et al. 1993).

## References

Ainslie, G. (2001) *Breakdown of will*. Cambridge University Press.

Ainslie, G. (2005) A précis of *Breakdown of will*. *Behavioral and Brain Sciences* 28(5), 635-673.

Bodner, R. & Prelec, D. (2001) The diagnostic value of actions in a self-signaling model. In: *Collected essays in psychology and economics*, ed. I. Brocas & J. Carillo Oxford University Press.

Davies, N. (1981) *Human sacrifice in history and today*. Morrow.

Frank, R. H., Gilovich, T. & Regan, D. (1993) Does studying economics inhibit cooperation? *Journal of Economic Perspectives* 7:159–71.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. & Rizzolatti, G. (1999) Cortical mechanisms of imitation. *Science* 286:2526–28.

Leach, C. W., Spears, R., Branscombe, N. R. & Doosje, B. (2003) Malicious pleasure: Schadenfreude at the suffering of another group. *Journal of Personality and Social Psychology* 84:932–43.

Weber, M. (1904/1958) *The protestant ethic and the spirit of capitalism*. Charles Scribners.