the games reflects daily life – that, as the authors put it, "experimental play often reflects patterns of interaction found in everyday life," and more specifically that variation in "market integration and the local importance of cooperation explain a substantial portion of the behavioral variation between groups" (sect. 1, para. 4).

Yet, data on the study populations published elsewhere reveal enough discrepancies to call these generalizations into question:

1. Among the Hadza, the only significant predictor of UG (ultimatum game) offers is camp size: the modal offer is 50% in large camps versus 20% in small camps (Marlowe 2004a, p. 179ff). Yet food sharing between households is greater in small camps than in large ones. So "generosity" in the UG is inversely related to real-life food sharing, something inconsistent with the target article conclusion but consistent with a signaling or reputation-based interpretation.

2. In the Conambo community, members of the Achuar coalition make much higher UG offers (mode = 50%) than do members of the Quichua coalition (25%). The target article ranks Achuar and Quichua as having identical "payoffs to cooperation" and very similar "aggregate market integration" (Fig. 5). The two coalitions do not differ in their frequency of between-household sharing of game, nor in the imbalance of such sharing (Patton 2004, p. 119). They do differ in the stability of the coalitions, which Patton argues explains the observed differences in UG offers, as well as the fact that among the Achuar (but not Quichua) higher status men share meat more widely (which Patton interprets as an alliance-building strategy). More recent data indicate that a reversal in coalition stability has led to a corresponding shift in relative generosity in game play, reinforcing this interpretation (Patton, personal communication).

3. Among the Tsimane, the five villages studied differ in closeness to market towns; although village membership is the best predictor of UG offers, the villages closer to town made smaller UG offers, contrary to the "market integration" hypothesis (Gurven 2004a). After analyzing a number of variables in relation to UG and PGG (public goods game) play, Gurven concludes that little "can be attributed to market exposure or acculturation, and the few differences that exist do not support the notion that exposure to modern markets produces game behavior similar to that found in the west" (Gurven 2004a, p. 217).

4. There is no relationship between generosity in the UG or the PGG and empirical data on food sharing between Ache individuals or households (Hill & Gurven 2004).

5. Neither the Tsimane (Gurven 2004a) nor the Ache (Hill & Gurven 2004) show any correlation between individual play in the UG versus the PGG.

6. The two New Guinea villages studied by Tracer (2004) differ substantially in degree of market integration, but UG offers do not differ significantly. Yet at the individual level, greater wealth, larger cash-crop gardens, and a history of wage labor, were all associated with higher UG offers.

These patterns suggest that generalizations regarding market integration and payoffs to cooperation are at best only part of the story; it is hard to see how the cross-cultural patterns claimed in the target article (which are based on subjective estimates of "market integration" and "payoffs to cooperation") could be very robust when so many of the within-society findings fail to find such patterns – even between different communities of the same culture.

Theoretical Implications. The "canonical model" of rational self-interest may still hold sway in much of economics and closely allied fields, but those of us with a more evolutionary approach to decision-making have a different take on this topic. First, "self-interest" is trumped by inclusive fitness considerations whenever the effects of one's actions alter the fitness of close relatives (Hamilton 1964) or well-defined sets of relatives even if not all are close (Jones 2000). Second (and more germane to the experimental findings), actions that sacrifice short-term self interest can be shown to be favored in various evolutionary regimes if they establish reputations (Hirshleifer 1987; Frank 2001; Milinski et al. 2002; Panchanathan & Boyd 2004), serve as commitment devices (Hirshleifer 2001; Irons 2001; Sosis & Alcorta 2003), or function as honest signals of underlying quality (Bliege Bird & Smith 2005; Boone 1998; Gintis et al. 2001). All of these strategic devices could be effective in attracting allies or discouraging competitors, thus recouping the short-term costs of providing public goods or engaging in costly enforcement of norms.

The hypothesis that reputation management is an important motivation for cooperation, generosity, and punishment of selfishness could account for at least some of the experimental results. For example, the target article notes that "the Tsimane and Machiguenga live in societies with little cooperation, sharing or exchange beyond the family unit. Ethnographically, both groups demonstrate little fear of social sanctions and seem to care little about local opinion" (sect. 8, para. 6). Indeed, several of the coauthors of the target article have proposed reputation-based explanations for the experimental results in the societies in which they work (see works cited above, as well as Alvard 2004; Gurven 2004b; Marlowe 2004b; Patton 2005; Tracer 2003).

However, other researchers cite experimental conditions (degree of anonymity, one-shot vs. repeated games) as evidence that reputation cannot explain the generosity exhibited in these and other studies (e.g., Bowles & Gintis 2003; Fehr & Henrich 2003; but see Haley & Fessler 2005 for a counterview). Because players do not maximize material payoffs from game play, alternative motivational structures (other-regarding, prosocial, strong reciprocity) are postulated to guide social interactions, and a history of group selection (cultural or genetic) is proposed to explain why these motivational structures exist (e.g., Bowles & Gintis 2003; Fehr & Fischbacher 2003; Gintis et al. 2003; Richerson et al. 2003).

I don't see how one can have it both ways: experimental conditions (e.g., anonymity, one-shot play, no opportunities for reputation management) strictly define the payoffs, but motivational structures or framing effects reflect daily life (market exposure, quotidian forms of collective action, etc.). This is the implicit argument in the target article's critique of the canonical model (or any self-interest based interpretation of the experimental results), yet it is internally inconsistent. If "individuals in experiments bring the preferences and beliefs that they have acquired in the real world into the decision-making situation" (sect. 9, para. 9), how is it that they cannot be guided by reputation issues because they are playing anonymously? I would like to see the authors' response address this discrepancy.

ACKNOWLEDGMENTS

I thank Michael Gurven, John Patton, and Richard Sosis for helpful comments.

Sociality and self interest

Vernon L. Smith

Departments of Economics and Law, George Mason University, Fairfax, VA 22030. **Vsmith2@gmu.edu**

Abstract: Selfishness narrowly defined as choosing dominant outcomes independent of context is widely rejected by experimentalists. Humans live in two worlds of personal and impersonal exchange; both are manifestations of human sociality, but the emphasis on preferences rather than cultural norms of personal exchange across time too much reflects a limited economic modeling, and fails to capitalize on the fresher experimental economics message of culture and diversity.

Economic theory is traditionally about proving theorems, not about studying economic behavior in a species whose largely autonomic sociality is as impressive as its intentional use of reason. The Scottish philosophers (e.g., David Hume, Adam Smith) understood that humans and their institutions achieve unintended ends, but somehow their ingenious discoveries got mislaid in our anthropocentric preoccupation with Cartesian logic (Smith 2003).

https://doi.org/10.1017/S0140525X05410146 Published online by Cambridge University Press

Humans live in two worlds – one of personal, the other of impersonal exchange; chimps, capuchins, and so on, live only in the former (de Waal 2005). Humans miraculously leveraged the former, the world of personal exchange, into cultures that enabled our ancestors to settle almost all the globe before they became herder-farmers.

Henrich et al.'s monumental and valuable cross-cultural study is *not* about human sociality – equity (equal opportunity), reciprocity, and sharing across time in personal exchange. It is about rejecting selfishness defined as choosing dominant outcomes independent of context, history, and the future. It is about attacking this narrow and misguided view of the self interest in social interactions, and shades of standard externality theory. It is also about adding arguments to utility functions – equity, reciprocity, and "fairness" as preferences, not as norms allowing self-betterment through gains from exchange across time.

I have argued (Smith 1998; 2003) that "property;" that is, human rights to act, arose endogenously by consent in social exchange; that norms like reciprocity – you have to give in order to receive – long preceded market exchange; and that these ancient emergent "laws" continue alive and well today, as revealed in the universal human expression, "I owe you one," in which A voluntarily acknowledges a debt to B for favors. Negative reciprocity is the endogenous policeman that reminds of debts not paid. It dictates punishment when the expectations of favors are not fulfilled.

These rules survive in the laboratory when subjects move in single-play trust games (TG) with or without punishment options, and cooperate with monotonous increasing average frequency in repeat play, as we vary the probability of being matched with the same person from zero to unity (McCabe et al. 1996). Also, when B sees what player A forgoes when risking defection by B, it doubles the cooperative response compared with B seeing that A has no alternative but to pass to B. This cannot be due to preferences regarding one's own and others' payoffs (McCabe et al. 2003). I see behavior in more austere Ultimatum Game (UG), Dictator Game (DG), and Public Good Games (PGG) as flatland projections of richer TG environments; the former are too easily over-interpreted in terms of utility rather than sociality. The evidence that children and young adults lean toward dominate outcome choices, and are vulnerable to drug abuse, suggests the likelihood that the superior benefits of delayed gratification from reciprocity have not yet become fully integrated into experience and behavior.

Our TG studies primed me for the central claim here that impersonal market exchange enhances, rather than crowds out, personal exchange as a viable hypothesis, credible if not predictable from extant theory. The cross-cultural data presented by Henrich et al. are consistent with this hypothesis, and the authors have done an excellent job of answering my earlier criticism that I believed their study suffered from a confounding of context and culture. As Henrich et al. note, the ad hoc need to adapt procedures to the field realities led to several variations, and these turn out to not materially affect the behavior. For future studies, I think they are now in a position to make such variations part of the design.

I believe such exercises in *deliberate* variation in the defined context are especially important in cross-cultural studies, in which there can be no presumption, given the wide variation in the techniques needed to convey comprehension, that different experimenters using different language, formats, descriptors, protocols, currencies, group sizes, and settings are all implementing the same "abstract game." Given all this variation, I don't know what it means to say the "researchers stuck to entirely abstract explanations of the game" (target article, sect. 4.3, para. 2). There is no such thing as a context-free experiment. All memory is autobiographical and context dependent, and our brains draw on that memory when deciding in an unfamiliar situation. What the experimenter thinks is "neutral," may not be so for the subjects. Hence the need to vary protocols and to empirically determine what is or is not a contextual "treatment." Recall that the canonical model predicts that none of these issues matter, yet they are obviously a consequence of variation in human sociality. When theory fails, it ceases to provide any guidelines as to how to understand "abstract" play, and you are back to the starting point.

These considerations are underlined by the wide variation found in our (Hoffman et al 1994) studies of UG – mean proposals varied from 27.8% to 44.1% with undergraduates. Our variations shifted distributions, without altering their basic shapes, which I think is an interesting and unique feature of the cross-cultural data. But how is this affected by deliberately varying protocols? I assume that field anthropologists have some conjectures as to what in a particular culture might account for these "outlying" results, and through deliberate design variations they can increase our understanding of the culture. For our North American culture, we explored "earned rights" and the buy-sell context, and we were surprised that we could move the data across such a range with these treatments and their conjunction. Also, we conjectured that the social connection – we called it "distance" – may not have gone far enough in the DG (Hoffman et al 1996b). We were surprised, not just that the "double blind" made a large economic difference – giving nothing rose from 20% to 64% – but that, by relaxing the double blind conditions incrementally, the results stepped down incrementally to the baseline.

For DG, the authors should consult Cherry et al. (2002), who compare endowments given by the experimenter with money earned by the subjects: the percent of dictators giving nothing jumps from 19% to 79%. And the latter rises to 97% when double blinded. Strengthening the sense that the stakes belong to the dictator, combined with "no one can know" his or her decision, all but eliminates dictator giving. We don't know how playing with the money you earned impacts UG, PGG, and cross-cultural results; it would be interesting to find out.

Methods do matter: Variation in experimental methodologies and results

Richard Sosis

Department of Sociology and Anthropology, The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem, 91905, Israel. **richard.sosis@uconn.edu**
www.anth.uconn.edu/faculty/sosis/

Abstract: Henrich et al.'s findings are generally consistent with the findings of other experimental field studies. However, the methodological variations deployed across field sites produced several systematic biases in their results. Hence, although the project is destined to become a watershed study, their results should be interpreted prudently.

Henrich and colleagues are to be commended for initiating what is surely a landmark interdisciplinary study that promises to have a profound impact on numerous fields in the social and biological sciences. Indeed, there is already a burgeoning anthropological and economic literature employing economic games among non-student populations (e.g., Carpenter et al. 2005; Lesorogol 2005b; Paciotti & Hadley 2003, Paciotti et al. 2005; Sosis & Ruffle 2003; 2004), some of which has been stimulated by Henrich's (2000) pioneering work among the Machiguenga. The findings of the target article are largely consistent with these studies: namely, non-student populations behave differently than students, the selfishness axiom is not supported, daily experiences influence game responses, and demographic factors have little impact on experimental decisions. The current study, as well as their companion co-edited volume, Henrich et al. (2004), will set the standard for the field. Nonetheless, I am concerned about the variance in methods employed across field sites and about how these variations influenced the experimental results.

Henrich et al. argue that the variation in their methodologies had little or no impact on the decision-making behavior of their subjects. Their data, however, cannot fully address this issue because there is no baseline data for any populations other than the Michigan and UCLA samples. We do not know whether method-

https://doi.org/10.1017/S0140525X05410146 Published online by Cambridge University Press