Word Counts:
Abstracts: 318
Main: 14902 (just main text, not captions, endnotes or tables)
References: 2542

'Economic Man' in Cross-cultural Perspective:

Behavioral Experiments in 15 Small-Scale Societies

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton and David Tracer

Affiliations: **Henrich**, Emory University, Dept. of Anthropology, Atlanta, GA 30322; **Boyd**, Dept. of Anthropology, University of California, Los Angeles CA 90095 rboyd@anthro.ucla.edu; **Bowles**, Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, New Mexico 87501, bowles@santafe.edu; **Camerer**, Div HSS 228-77, Caltech, Pasadena CA 91125 USA, camerer@hss.caltech.edu; **Fehr**, University of Zurich, Bluemlisalpstrasse 10, CH-8006, Zurich, Switzerland; **Gintis**, Santa Fe Institute hgintis@comcast.net; **McElreath**, University of California Davis, Dept. of Anthropology, One Shields Ave, Davis, CA, mcelreath@ucdavis.edu; **Alvard**, Dept. of Anthropology, 4352 TAMU, Texas A&M University, College Station, Texas 77843-4352, Alvard@tamu.edu; **Barr**, Centre for the Study of African Economies, University of Oxford, St. Cross Building, Manor Road, Oxford OX1 3UL, abigail.barr@economics.ox.ac.uk; **Hill**, Dept. of Anthropology, University of New Mexico, kimhill2@unm.edu; **Gurven**, Dept. of Anthropology, University of California-Santa Barbara, Santa Barbara, CA 93106, gurven@anth.ucsb.edu; **Patton**, Dept. of Anthropology, Washington State University, Pullman, WA 99164-4910, pattonj@mail.wsu.edu; **Tracer**, Dept. of Anthropology &, Health & Behavioral Sciences, University of Colorado at Denver, Denver, CO 80217, dtracer@carbon.cudenver.edu; **Ensminger**, Div HSS 228-77, Caltech, 1200 East California Blvd., Pasadena, CA 91125-7700, jensming@hss.caltech.edu; **Gil-White**, Dept. of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia PA 19104-6196, fjgil@psych.upenn.edu; **Marlowe**, Dept. of Anthropology, Peabody Museum, 11 Divinity Avenue, Harvard University, Cambridge, Massachusetts 02138, fmarlowe@fas.harvard.edu; **Henrich**, ORC Macro, Natalie.J.Henrich@orcmacro.com

Abstracts

Short Abstract

While existing experimental work has revealed consistent deviations from self-interest, it cannot determine whether this results from universal human motives or is a property of the university student subjects. To address this, we undertook a cross-cultural experimental study in fifteen small-scale societies, and found that (1) the self-interest model fails in all of the societies studied, (2) there is more variability than has previously been observed, (3) group-level differences explain a substantial portion of the behavioral variation across societies, (4) the available individual-level economic and demographic variables do not robustly explain behavior across different societies, and (5) experimental play often reflects the patterns of everyday life.

Long Abstract

Researchers from across the social sciences have found consistent deviations from the predictions of the canonical model of self-interest in hundreds of experiments from around the world. This research, however, cannot determine whether this uniformity results from universal patterns of human behavior or from the limited cultural variation available among the university students used in virtually all prior experimental work. To address this, we undertook a cross-cultural study of behavior in Ultimatum, Public Goods, and Dictator Games in fifteen small-scale societies exhibiting a wide variety of economic and cultural conditions.

We found, first, that the canonical model—based on pure self-interest—fails in all of the societies studied. Second, the data reveals substantially more behavioral variability across social groups than has been found in previous research. Third, group-level differences in economic organization and the structure of social interactions explain a substantial portion of the behavioral variation across societies: the higher the degree of market integration and the higher the payoffs to cooperation in everyday life, the greater the level of prosociality expressed in experimental games. Fourth, the available individual-level economic and demographic variables do not robustly explain game behavior, either within or across groups. Fifth, in many cases experimental play appears to reflect the common interactional patterns of everyday life.

Keywords: Self-interest, altruism, cooperation, Ultimatum Game, Public Goods Game, cross-cultural research, experimental economics, game theory

Since “Selfishness examined…” appeared in these pages, more than fifteen years ago, many additional experiments have strongly confirmed the doubts expressed by Caporael and her collaborators (1989) concerning the adequacy of self-interest as a behavioral foundation for the social sciences. Experimental economists and others have uncovered large, consistent deviations from the textbook predictions of *Homo economicus* (Camerer, 2003; Fehr, Fischbacher, & Gaechter, 2002; Hoffman, McCabe, & Smith, 1998; Roth, 1995). Hundreds of experiments in dozens of countries, using a variety of game structures and experimental protocols, suggest that in addition to their own material payoffs, students care about fairness and reciprocity and will sacrifice their own gains to change the distribution of material outcomes among others, sometimes rewarding those who act prosocially and punishing those who do not. Initial skepticism about such experimental evidence has waned as subsequent work involving high stakes and ample opportunity for learning has repeatedly failed to modify these fundamental conclusions.

This multitude of diverse experiments creates a powerful empirical challenge to what we call the *selfishness axiom*—the assumption that individuals seek to maximize their own material gains in these interactions and expect others to do the same.[1] However, key questions remain unanswered. Do such consistent violations of the canonical model provide evidence of universal patterns that characterize our species? Or, do individuals’ economic and social environments shape their behavior, motivations and preferences? If so, are there boundaries on the malleability of human nature, and which economic and social conditions are most involved? Are there cultures that approximate the canonical account of purely self-regarding behavior? Are inclinations towards fairness (equity) and ‘tastes’ for punishing unfairness better explained statistically by individuals' attributes such as their sex, age, education and relative wealth, or by the attributes of the individuals' group?

Existing research cannot answer such questions because virtually all subjects have been university students. While there are modest differences among student populations throughout the world (Roth, Prasnikar, Okuno-Fujiwara, & Zamir, 1991), these differences in subjects and settings are small compared to the full range of human social and cultural environments. To broaden this inquiry, we undertook a large cross-cultural study using Ultimatum, Public Goods, and Dictator Games. Twelve experienced field researchers, working in twelve countries on four continents and New Guinea, recruited subjects from fifteen small-scale societies exhibiting a wide variety of economic and social conditions. Our sample of societies consists of three groups of foragers, six slash-and-burn horticulturalists, four nomadic herders, and two groups of small-scale agriculturalists.

Our overall results can be summarized in five major points: first, there is no society in which experimental behavior is fully consistent with the selfishness axiom; second, there is much more variation between groups than previously observed, although the range and patterns in the behavior indicate that there are certain constraints on the plasticity of human sociality; third, differences between societies in market integration and the local importance of cooperation explain a substantial portion of the behavioral variation *between* groups; fourth, individual-level economic and demographic variables do not consistently explain behavior within or across groups; fifth, experimental play often reflects patterns of interaction found in everyday life. Below, we describe the experimental methods used and give a comparative overview of the societies studied. We then present and interpret our findings. More extensive details about each society, results and methods can be found in Henrich *et. al.* (2004).

(1) Experimental Games and Behavior in Student Populations

The three experiments we deployed, the Ultimatum (UG), Dictator (DG), and Public Goods Games (PGG), have been extensively studied in complex market societies, where subjects have mostly been college students. Here, we lay out the basic games, and briefly summarize the typical findings in student populations. For extensive reviews see Kagel & Roth (1995) and Camerer (2003).

The Ultimatum and Dictator Games

The UG is a simple bargaining game that has been extensively studied. In this game, subjects are paired, and the first player, often called the "proposer," is provisionally allotted a divisible "pie" (usually money). The proposer then offers a portion of the total pie to a second person, called the "responder." The responder, knowing both the offer and the amount of the pie, can then either accept or reject the proposer's offer. If the responder accepts, he receives the offer and the proposer gets the remainder (the pie minus the offer). If the responder rejects the offer, then neither receives anything. In either case, the game ends; the two subjects receive their winnings and depart. Players are typically paid in cash and are anonymous to other players, but not to the experimenters (although experimentalists have manipulated these variables). In all our experiments, players were anonymous to each other, and the games used substantial sums of money (in the appropriate currency). For this game, the canonical model (i.e., all participants maximize their income and this is known by all of them) predicts that responders, faced with a choice between zero and a positive payoff, should accept any positive offer. Knowing this, proposers should offer the smallest non-zero amount possible. In every experiment yet conducted the vast majority of participants violated this prediction. The DG is the same as the UG, except that responders are not given an opportunity to reject—they simply get whatever the proposer dictates.

In student populations, modal offers in the UG are almost always 50%, and mean offers are between 40 and 45%. Responders reject offers of 20% about half the time, and rejection is associated with emotional activation in the insula cortex (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). In the DG, modal offers are typically 0% and means usually fall in the 20 to 30% range, although DG results are more variable than in the UG.

The Public Goods Game

The PGG shows how people behave when individual and group-interests conflict. We used two variants: the 'Voluntary Contributions' (VC) and the 'Common-Pool Resources' (CPR) formats. In the VC version, players receive some initial monetary endowment, and then have the simultaneous opportunity to anonymously contribute any portion of their endowment (from zero to the full endowment) to the group fund. Whatever money is in the group fund after all players have contributed is augmented by 50% (or sometimes doubled), and then distributed equally among all players regardless of their contribution. The payoff structure of the CPR format is identical, except that instead of receiving an endowment, players can make limited withdrawals from the group fund. Whatever remains in the fund (the common pool) after everyone has withdrawn is increased by 50% or doubled, and distributed equally among all group members. The game is not repeated. Selfish subjects may calculate that, independent of the actions taken by the other players, contributing as little (in the VC) or withdrawing as much as possible (in the CPR) maximizes their monetary payoffs: Free-riding is thus the dominant strategy for selfish subjects.

Students in one-shot Public Goods Games contribute a mean amount between 40% and 60%, although there is a wide variance, with most contributing either everything or nothing

(Henrich & Smith, 2004; Ledyard, 1995; Sally, 1995). While this is fairly robust, people are sensitive to the costs of cooperation and repeated play. Raising the amount by which the augmentation percentage of the common pool produces an increase in contributions (Andreoni & Miller, 2002). When the PGG is played repeatedly with the same partners, the level of contribution declines towards zero, culminating in most subjects refusing to contribute to the common pool (Andreoni, 1988; Fehr & Gächter, 2000, 2002).

The two major concerns with interpreting experimental data—stake size and familiarity with the experimental context—have largely been put to rest. For stakes, some have argued that as the stakes increase, the "cost of being non-selfish" increases, so selfish behavior should increase. Were this true, it would show that in behaving unselfishly, people respond to costs and benefits (as they do in many games; e.g., Ledyard, 1995 and Miller and Andreoni, 2002). But evidence of responding to the 'cost of being non-selfish' does not suggest that unselfish behavior is unimportant or extinguished at high stakes. Indeed, in the UG, raising the stakes to quite high levels (e.g., three months' income) does not substantially alter the basic results (Camerer & Hogarth, 1999; Cameron, 1999; Hoffman, McCabe, & Smith, 1996; List & Cherry, 2000; Slonim & Roth, 1998). In fact, at high stakes, proposers tend to offer a little more, and responders remain willing to reject offers that represent small fractions of the pie (e.g., 20%) even when the pie is large (e.g., $400 in the U.S., see List & Cherry, 2000). Similarly, the results do not appear to be due to a lack of familiarity with the experimental context. Subjects do not change their behavior in any systematic way when they participate in several replications of the identical experiment (Fehr & Gächter, 2002; Knez & Camerer, 1995; List & Cherry, 2000).

Several researchers have tested the effects of demographic variables on behavior in experimental games (Camerer, 2003). The general result is that demographic effects are nonexistent, inconsistent and/or weak. In the UG, female students reject somewhat less often, but no differences emerge for offers. In the DG, no gender differences have been found. Similarly, the age of adult subjects has not turned up as an important predictor in any of our games, or among the handful of results from non-student populations in the U.S. (Carpenter, Burks, & Verhoogen, 2002; Smith, 2003). Thus, our cross-cultural results are consistent with existing findings on demographic variables. However, there is intriguing evidence that children behave selfishly, but gradually behave more fair-mindedly as they grow older, up to age 22 or so (Harbaugh & Krause, 2000; Harbaugh, Krause, & Liday, 2002; Murnighan & Saxon, 1998). An important exception is that about a third of autistic children and adults offer nothing in the UG (Hill & Sally, 2004); presumably their inability to imagine the reactions of responders leads them to behave, ironically, in accordance with the canonical model.

Behavioral economists have been remarkably successful in explaining the experimental behavioral of students by adding social preferences (especially those related to equity, reciprocity, and fairness) to game theoretical models (Camerer, 2003; Fehr & Schmidt, 1998). Our endeavor aims at the foundation of these proximate models by exploring the nature of non-selfish preferences.

(2) The Cross-cultural Behavioral Experiments Project

Early cross-cultural economic experiments (Cameron, 1999; Roth et al., 1991) showed little variation among university students. However, in 1996 a surprising finding broke the consensus: the Machiguenga, slash-and-burn horticulturalists living in the southeastern Peruvian Amazon, behaved much less prosocially than student populations around the world (Henrich, 2000). The UG "Machiguenga outlier" sparked curiosity among a group of behavioral scientists: was this simply an odd result, perhaps due to the unusual circumstances of the experiment, or had Henrich tapped real behavioral differences, perhaps reflecting the distinct economic circumstances or cultural environment of this Amazonian society? In November 1997, the MacArthur Foundation Research Network on the Nature and Origin of Preferences brought 12 experienced field workers and several behavioral economists together in a three-day meeting at UCLA. During this meeting we redesigned the experiments—typically conducted in computer labs at universities—for field implementation in remote areas among non-literate subjects. Two years later, once all of our team had returned from the field, we reconvened to present, compare and discuss our findings. Here we summarize these findings thus far (however, Phase II is underway).

The Experiments

Overall, we performed 15 Ultimatum, six Public Goods, and three Dictator Games, as well as two control experiments in the U.S., at UCLA and the University of Michigan. All of our games were played anonymously, in one-shot interactions, and for substantial real stakes (the local equivalent of one or more days' wages). Because the UG was administered everywhere (n = 564 pairs), we will concentrate on these findings and their implications, and make only some references to our other games (see Henrich et al., 2004).

Ethnographic Description

Figure 1 shows the location of each field site, and Table 1 provides some comparative information about the societies discussed here. In selecting these, we included societies both sufficiently similar to the Machiguenga to offer the possibility of replicating the original Machiguenga results, and sufficiently different from one another to provide enough social, cultural and economic diversity to allow an exploration of the extent to which behaviors covary with local differences in the structures of social interaction, forms of livelihood, and other aspects of daily life.

Figure 1. Locations of small-scale societies

In Table 1, the column 'Economic Base' provides a general classification of the production system of each society. *Horticulturalists* rely primarily on slash-and-burn agriculture, which involves clearing, burning and planting gardens every couple of years. All the horticulturalists included here also rely on some combination of hunting, fishing and gathering. We have classified the Aché's economic base as *Horticulture–Foraging* because they were full-time foragers until about three decades ago, and still periodically go on multi-week foraging treks, but have spent much of the last few decades as manioc-based horticulturalists. The Au and Gnau of Papua New Guinea are classified as *Foraging/Horticulture* because, despite planting small swidden gardens, they rely heavily on harvesting wild sago palms for calories, and game for protein. Unlike foragers and horticulturalists, *Pastoralists* rely primarily on herding. *Agro-*

pastoralists rely on both small-scale sedentary agriculture and herding. We labeled the Orma, Mongols and Kazakhs as pastoralists because many people in these societies rely completely on herding, although some members of all three groups do some agriculture. The Sangu are labeled *Agro-pastoralists* because many people rely heavily on growing corn, while others rely entirely on animal husbandry (consequently, we sometimes separate Sangu herders and Sangu farmers).

The column 'Residence' classifies societies according to the nature and frequency of their movement. *Nomadic* groups move frequently, spending as few as a couple of days in a single location, and as long as a few months. *Semi-nomadic* groups move less frequently, often staying in the same location for a few years. Horticultural groups are often semi-nomadic, moving along after a couple of years in search of more abundant game, fish, wild foods and fertile soils. *Transhumant* herders move livestock between two or more locales in a fixed pattern, often following the good pasture or responding to seasonal rainfall patterns. *Bilocal* indicates that families maintain two residences, and spend part of the year at each residence. The Machiguenga, for example, spend the dry season living in villages along major rivers, but pass the wet season in their garden houses, which may be located three or more hours from the village. The *Bilocal–Semi-nomadic* classification indicates that the Machiguenga, for example, were traditionally semi-nomadic, but have more recently adopted a bilocal residence pattern. Similarly, the Aché are classified as *Sedentary–Nomadic* because of their recent transition from nomadic foragers to sedentary horticulturalists.

Table 1. Ethnographic Summary of Societies

Group	Language Family	Environment	Economic Base	Residence	Complexity	Researcher	PC	AMI
Machiguenga	Arawakan	Tropical Forest	Horticulture	Bilocal semi nomadic	Family	Henrich, Smith	1	4.5
Quichua	Quichua	Tropical Forest	Horticulture	Sedentary/ Semi-nomadic	Family	Patton	1	2
Achuar	Jivaroan	Tropical Forest	Horticulture	Sedentary/ Semi-nomadic	Family plus extended ties	Patton	5	2.5
Hadza	Khoisan/Isolate	Savanna-Woodlands	Foraging	Nomadic	Band	Marlowe	4	1.25
Aché	Tupi-Guarani	Semi-tropical Woodlands	Horticulture/ Foraging	Sedentary-Nomadic	Band	Hill, Gurven	6	5
Tsimane	Macro-Panoan Isolate	Tropical Forest	Horticulture	Semi-nomadic	Family	Gurven	1	2.75
Au	Torricelli/ Wapei	Mountainous Tropical Forest	Foraging/ Horticulture	Sedentary	Village	Tracer	3	4.75
Gnau	Torricelli/ Wapei	Mountainous Tropical Forest	Foraging/ Horticulture	Sedentary	Village	Tracer	3	5
Mapuche	Isolate	Temperate Plains	Small scale farming	Sedentary	Family plus extended ties	Henrich	2	4
Torguuds	Mongolian	High latitude desert Seasonally-flooded grassland	Pastoralism	Transhumance	Clan	Gil-White	2	9
Kazakhs	Turkic	High-latitude Desert Seasonally-flooded grassland	Pastoralism	Transhumance	Clan	Gil-White	2	9.25
Sangu (farm/herd)	Bantu	Savanna-Woodlands Seasonally-flooded grassland	Agro-Pastoralists	Sedentary or Nomadic	Clan-Chiefdom	McElreath	5	6.5 6.75
Orma	Cushitic	Savanna-Woodlands	Pastoralism	Sedentary or Nomadic	Multi-Clan Chiefdom	Ensminger	2	9.25
Lamalera	Malayo-Polynesian	Island Tropical coast	Foraging-Trade	Sedentary	Village	Alvard	7	9
Shona	Niger-Congo	Savanna-Woodlands	farming	Sedentary	Village	Barr	5	10

The column 'Language Family' provides the current linguistic classification for the language traditionally spoken by these societies, and is useful because linguistic affinity provides a rough measure of the cultural relatedness of two groups. The classification of the Mapuche, Hadza, Tsimane and New Guinean languages demand comment. There is no general agreement about how to classify *Mapudungun* (the Mapuche's language) with the other language groups of South America. Similarly, although Hadza was once considered a Khoisan language, distantly related to the San languages of southern Africa, agreement about this is diminishing. The Tsimane language resembles Moseten (a Bolivian group similar to the Tsimane), but otherwise these two seem unrelated to other South American languages, except perhaps distantly to Panoan. Finally, because of the linguistic diversity found in New Guinea, we have included both the language phylum for the Au and Gnau, Torricelli, and their language family, Wapei.

The column, 'Complexity' refers to the anthropological classification of societies according to their political economy (Johnson & Earle 2000). *Family-level* societies consist of economically-independent families that lack any stable governing institutions or organizational decision making structures beyond the family. Societies classified as *Family plus extended ties* are similar to family-level societies, except that such groups also use extended kin ties or non-kin alliances, for specific purposes such as warfare. In these circumstances, decision making power remains ephemeral and usually diffuse. *Bands* are composed of both related and unrelated families who routinely cooperate in economic endeavors. Decision making relies substantially on group consensus, although the opinions of prestigious males often carry substantial weight. *Villages and Clans* are both corporate groups of the same level of complexity, and both are typically larger than bands. *Clans* are organized around kinship, tracked by lineal descent from a common ancestor. Decision-making power is often assigned according to lineage position, but achieved status plays some role. *Villages* operate on the same scale of social and political organization as clans, but usually consist of several unrelated extended families. Decision making is often in the hands of a small cadre of older, high status men. At a larger scale of organization, *Multi-Clan Corporate groups* are composed of several linked clans, and are governed by a council of older high status men—assignment to such councils is often jointly determined by lineal descent, age, and achieved prestige. Multi-clan corporations sometimes act only to organize large groups in times of war or conflict, and may or may not play an important economic role. Often larger than multi-clan corporations, *Chiefdoms* are ruled by a single individual or family and contain several ranked clans or villages. Both individual ranks and that of clans/villages usually depend on real or customary blood relations to the chief. Political integration and economic organization in chiefdoms is more intense than in multi-clan corporate groups, and chiefs often require subjects to pay taxes or tribute.

The two remaining columns on Table 1, aggregate market integration (AMI) and payoffs to cooperation (PC), refer to rankings we constructed on the basis of ethnographic investigations; we explain these below.

(3) Experimental Results

Substantial cross-cultural variability

The variability in Ultimatum Game behavior across the groups in our study is larger than that previously observed in large-scale, industrialized societies (e.g., Camerer, 2003, chapter 2). Prior work comparing UG behavior among university students from Pittsburgh, Ljubljana (Slovenia), Jerusalem, Tokyo (Roth et al., 1991) and Yogyakarta (Indonesia: Cameron, 1999) revealed little group variation. In contrast, Figure 2 summarizes our UG results from 15 different small-scale societies and shows substantial variation. While mean UG offers in standard experiments in industrialized societies are typically between 40% and 50% (see Camerer 2003, Table 2.2), the mean offers from proposers in our sample span a range from 26% to 58%—both below and above the typical behavior (Figure 2 and Table 2 presents additional details). While modal UG offers are consistently 50% among university students, our sample modes vary from 15% to 50%, although the 50/50 offer is clearly popular in many groups. As a student benchmark, we have included UG data from Roth et. al.'s (1991) Pittsburgh study.[2]

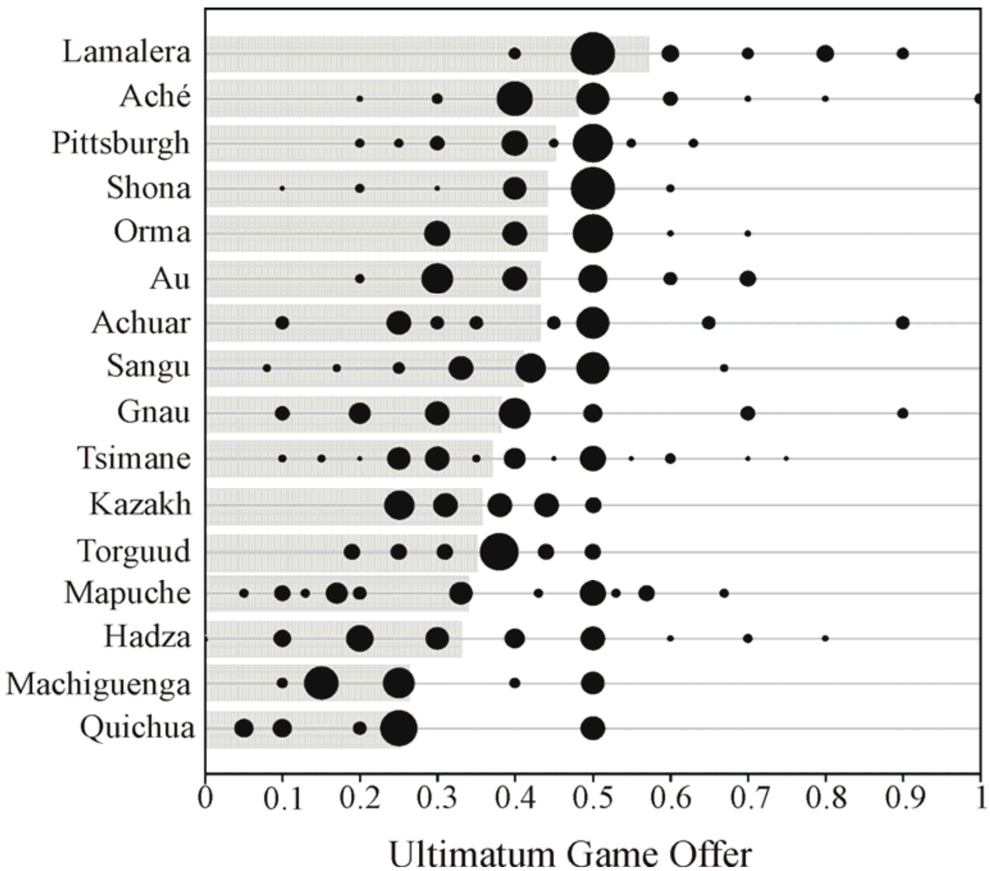

Figure 2. A bubble plot showing the distribution of UG offers for each group. The size of the bubble at each location along each row represents the proportion of the sample that made a particular offer. The right edge of the lightly shaded horizontal gray bar gives the mean offer for that group. Looking across the Machiguenga row, for example, the mode is 0.15, the secondary mode is 0.25, and the mean is 0.26.

Table 2 : Ultimatum Game Experiment Summary Statistics

Group	# Pairs	% female	Stake	Mean	Mode (% sample)[3]	Rejections	Low rejections[4]
Lamalera[5]	19	55	10	0.57	0.50 (63%)	4/20 (sham)[6]	3/8 (sham)
Aché	51		1	0.48	0.40 (22%)	0/51	0/2
Shona (resettled)	86	45	1	0.45	0.50 (69%)	6/86	4/7
Shona (all)	117	46	1	0.44	0.50 (65%)	9/118	6/13
Orma	56	38	1	0.44	0.50 (54%)	2/56	0/0
Au	30	48	1.4	0.43	0.3 (33%)	8/30	1/1
Achuar	14	50	1	0.43	0.50 (36%)	2/15[7]	1/3
Sangu (herders)	20	50	1	0.42	0.50 (40%)	1/20	1/1
Sangu (farmers)	20	50	1	0.41	0.50 (35%)	5/20	1/1
Sangu	40	50	1	.41	0.50 (38%)	6/40	2/2
Shona (unresettled)	31	48	1	0.41	0.50 (55%)	3/31	2/6
Hadza (big camp)	26	50	3	0.40	0.50 (35%)	5/26	4/5
Gnau	25	46	1.4	0.38	0.4 (32%)	10/25	3/6
Tsimane	70	51	1.2	0.37	0.5/0.3	0/70	0/5
Kazakh	10	45	8	0.36	0.38 (50%)	0/10	0/1
Torguud	10	50	8	0.35	0.25 (30%)	1/10	0/0
Mapuche	31	13	1	0.34	0.50/0.33	2/31	2/12
Hadza (all camps)	55	50	3	0.33	0.20/0.50	13/55	9/21
Hadza (small camp)	29	51	3	0.27	0.20 (38%)	8/29	5/16
Quichua	15	48	1	0.25	0.25 (47%)	0/14[6]	0/3
Machiguenga	21	19	2.3	0.26	0.15/0.25	1	1/10

On the responder side of the UG (Figure 3), rejection rates are also quite variable. In some groups, rejections are extremely rare, even in the presence of low offers, while in others, rejection rates are substantial, including frequent rejections of offers *above* 50%. Among the Kazakh, Quichua, Aché and Tsimane, we observe zero rejections out of 10, 14, 51 and 70 proposer offers, respectively. And, while offers among the Aché were mostly at or near 50%, 57% of the offers to Quichua and 47% of offers to Tsimane were at or below 30%—yet all were accepted. Similarly, Machiguenga responders rejected only one offer, despite the fact that over 75% of their offers were below 30% of the pie. At the other end of the rejection scale, Hadza rejected 24% of all offers and 43% (9/21) of offers 20% and below. Unlike the Hadza and other groups who preferentially rejected low offers, the Au and Gnau of Papua New Guinea rejected offers both below *and* above 50%, with nearly equal frequency. University student responders fall towards the upper end of the rejection scale (with more rejections than average), but still reject less than groups like the Au, Gnau, Sangu farmers, and Hadza, all of whom rejected positive offers with greater frequency than, for example, the Pittsburgh subjects in Roth et al. (1991).

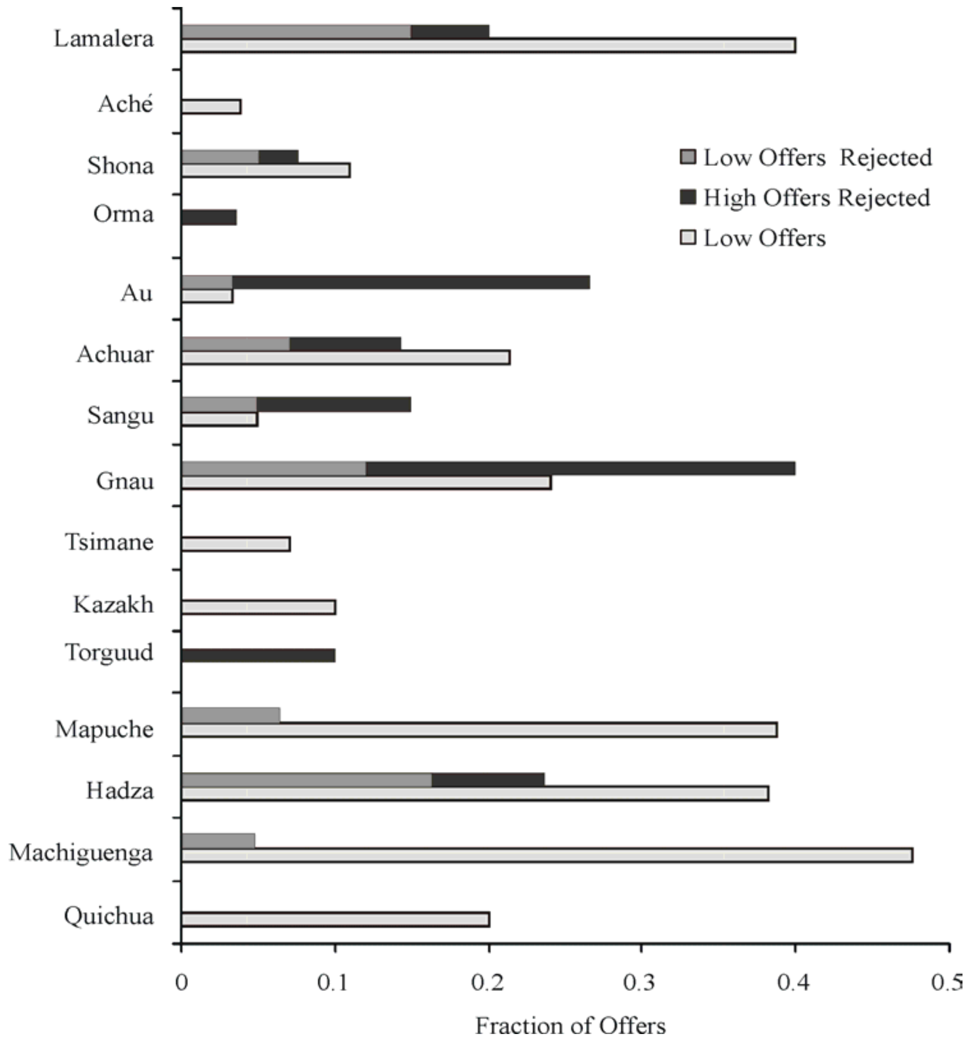

Figure 3. Summary of Responder's Behavior. The lightly shaded bar represents the fraction of offers that were less than 20% of the pie. The length of the darker shaded bar gives the fraction of all Ultimatum Game offers that were rejected, and the gray part of the darker shaded bar gives the number of these low offers that were rejected as a fraction of all offers. The low offers plotted for the Lamalera were sham offers created by the investigator.

Table 3: Summary of Public Good Experiments

Group	Format	Grp. size	MPCR[8]	Samp. Size	Stake[9]	Mean	Mode[10]	% Full Coop	% Full Defect
Michigan[11]	CPR	4	0.375	64	0.58	0.43	0 (33%)	26	33
Machiguenga	CPR	4	0.375	21	0.58	0.22	0 (38%)	0	38
Tsimane	VC	4	0.50	134	0.75	0.54	0.67 (17%)	1.5	5
Mapuche[12]	VC	5	0.40	12	0.33	0.34	0.1 (42%)	0	0
Huinca	VC	5	0.40	12	0.33	0.58	0.5 (25%)	17	0
Ache	VC	5	0.40	64	1	0.65	0.40 (30%)	3.1	1.6
Orma	VC	4	0.50	24	0.5	0.58	0.40 (37%)	25	0

As in the UG, our data from Public Goods Games, which include both VC and CPR versions, show much greater variation than previous experiments in industrialized societies. Typical distributions of PGG contributions from university students have a 'U-shape' with the mode at full defection (zero given to the group) and a secondary mode at full cooperation (everything to the group). While the format of the games does impact the results (e.g., people tend to give more in CPR vs. VC versions), the mean contributions still usually end up between 40% and 60%. Table 3 shows that our cross-cultural data provide some interesting contrasts with this pattern. The Machiguenga, for example, have a mode at full defection, but lack any fully cooperative contributions, which yields a mean contribution of 22%. By direct comparison (the protocol and experimenters were identical to the Machiguenga experiment), students at the University of Michigan produced the typical bimodal distribution, yielding a mean contribution of 43%. While both the Aché and Tsimane experiments yielded means similar to those found in industrialized societies, the shape of their distributions could not have been more different: they have unimodal, not bimodal, distributions. Their distributions resemble *inverted* American distributions with few or no contributions at zero or 100%. Like the Aché and Tsimane, the Huinca and Orma show modes near the center of the distribution, at 40% and 50% respectively, but they also have secondary peaks at full cooperation (100%)—and *no* contributions at full defection.

Violations of the Selfishness Axiom

The selfishness axiom was violated in some way in every society we studied, across all three different experimental games (DG, UG and PGG). Focusing on the UG, either proposer or responder behavior, or both, violated the axiom. Yet, responder behavior was consistent with selfish motives in several groups, unlike typical university students. As shown in Table 2, responders from the Aché, Tsimane, Machiguenga, Quichua, Orma, Sangu herders, and Kazakhs all have rejection rates of less than 5%, roughly consistent with the canonical model. For some groups these low rejection rates are not informative because all the offers were near 50/50 (e.g. the Aché and Sangu), so no one in the group received low offers. However, proposers in several societies made numerous low offers that were not rejected. The selfishness axiom accurately predicts responder behavior for about half of our societies, even though it generally fails to predict the responder behavior of university students. Like university students, Au, Gnau, Sangu farmers, and Hadza subjects rejected positive offers and thereby violated the axiom.

The issue of whether proposer behavior is consistent with the selfishness axiom is more complicated. Table 2 or Figure 2 shows that proposers are not making offers consistent with the standard game theoretical prediction based on the selfishness axiom, which requires that proposers offer the smallest positive amount—because they believe that the responders are seeking to maximize only their income from the game. In none of our societies was this behavior common.

Perhaps, however, proposers' behavior can be understood as income maximizing *given the presence* of responders willing to reject low offers. Among university subjects, it is generally thought that offers are fairly consistent with expected income-maximizing strategies *given* the empirical distribution of actual rejections across offers (Roth et al., 1991). Our results and analyses suggest that this is unlikely to be the case in several of the groups studied. For the groups in which at least one offer was rejected, we used the responder data to estimate an income-maximizing offer (IMO), and then compared this estimate to the group's mean offers. Intuitively, the IMO is the offer that an income-maximizing proposer would make assuming he knows the distribution of what responders in his group will accept (and is neutral toward economic risk, an important qualification we will return to shortly).

Figure 4 compares the actual mean offers from proposers (on the y-axis) with their corresponding IMO's (calculated from responder data, on the *x*-axis) for the various societies. The mean offer/IMO pairs for each society are plotted as points next to the societies' names. Look first at the midpoint and ignore the ellipses around them. *Every* group is above the unity line where mean UG offer = IMO. This unity line is where the average offer would lie if the average offer in each group were perfectly calibrated to that group's empirical IMO. When the mean UG offer is *above* the unity line, proposers are being '*generous*' given the likelihood of rejection at each offer level (i.e., they are offering more than selfishness alone would motivate them to offer).

To assess the statistical significance of how far mean offers depart from the estimated IMO, each point in Figures 4a-b is surrounded by an elliptical two-dimensional (2D) 90% "confidence interval".[13] A one-dimensional 90% confidence interval is a range of numbers that has a 90% chance of containing the true value of the statistic of interest. A 2D interval is the same idea extended to a *pair* of statistics. Using a statistical method called "bootstrapping," we can use the data we have to judge how differently the results might have turned out if the experiment had been done (hypothetically) over and over. The interval of bootstrapped values that results enables us to judge how confident we can be that the mean offer would almost always be above the IMO if our experiments were repeated.

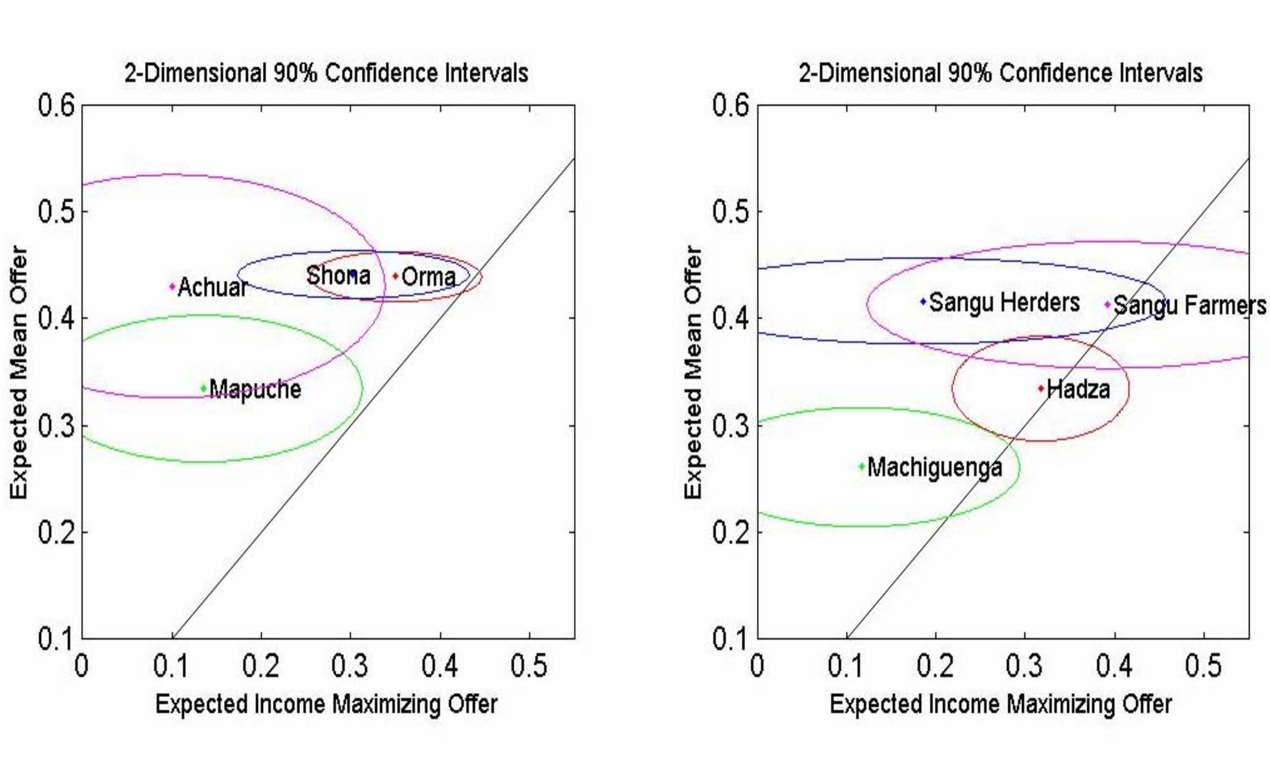

Figure 4. Two-dimensional 90% confidence intervals of the mean UG offers in various groups plotted against the expected income maximizing offers (estimated from observed distribution of rejections). Intervals show loci of possible mean offers and expected IMO's randomly resampled (bootstrapped) from samples. We were unable to estimate the IMO for societies with no rejections (Quichua, Tsimane, Ache, Kazakhs), or societies in which rejections bore no systematic relationship to offers (Au, Gnau).

Now return to the question of whether the average offer is above the IMO— i.e., did proposers offer significantly more than they had to, to maximize their earnings (given that some responders rejected low offers)? That question is answered at a glance, for a particular group, by simply observing whether the entire 2D ellipse for that group lies above and left of the 45-degree unity line. The two graphs plot separately those societies in which we can be quite confident the mean offer is clearly above the IMO (Figure 4a), and those for which we cannot be fully confident the mean offer is truly above the IMO (Figure 4b). Roughly half the societies clearly lie to the upper left, with their mean offers above their IMOs. The others also lie in upper left, but we cannot be too confident that their means are above their IMOs, *although* the ellipses only slightly overlap the 45-degree unity line for the Machiguenga and the Sangu herders.[14]

It is possible that such high offers are consistent with a more conventional extension of the selfishness axiom—aversion toward taking a chance on either getting a high or a low money payoff ("risk aversion" in economic language). It is a common (though not universal) observation that people prefer a certain amount of money to a gamble with the same expected payoff. Economists model this behavior by assuming that people seek to maximize their expected utility, and that utility is a concave function of income (diminishing psychophysical returns—earning an extra dollar is worth less in utility terms on top of a lot of other dollars, compare to a smaller number of dollars).

For example, suppose a subject estimates that an offer of 40% of the pie will be accepted for sure (leaving 60% for the proposer), and that an offer of 10% will be accepted with probability 2/3. If she were risk averse, she might value the certainty of keeping 60% of the pie more than the 2/3 chance of keeping 90% (and a 1/3 chance of getting nothing). In this case the expected monetary gain is the same for the two offers (namely, 60% of the pie), but the *expected utility* of the certain outcome is greater. Thus, a highly risk averse subject might make a high offer even if the probability of rejection of a low offer were small.

To explore whether risk aversion can explain the fact that average offers are so much higher than IMO's in most of our samples, we measured the degree of risk-aversion both indirectly and directly. The indirect measurement asks what degree of aversion toward risk is necessary to make the risk-adjusted IMO equal to the mean offer. To answer this we transformed the game payoffs into utilities, by assuming that the utility function for money is a power function x^ρ of the money amount x. ρ is a standard measure of the degree of risk aversion. For each group we estimated the value of ρ that would make the observed mean offer a utility maximizing offer given the distribution of actual rejection frequencies.[15]

As already noted in Figure 4b, the Hadza and the Sangu farmers were approximately expected income maximizers—that is, their average offers are consistent with expected utility maximization for risk neutral individuals. But for the other groups—Orma, Sangu herders, Machiguenga, Mapuche, and Shona—the implied levels of risk aversion are implausibly high. Even for the least extreme case, the Shona, the degree of risk aversion necessary to make their behavior consistent with expected utility maximization implies that they would be indifferent between an even chance that an offer of 1 out of 10 dollars would be accepted (an expected payoff of $4.5) and getting only $.04 for sure.[16] Clearly, an individual with this degree of risk aversion would be unable to function in uncertain environments.

Risk aversion was also measured *directly* among the Mapuche and the Sangu by offering subjects a series of risky choices between gambles with different probabilities of monetary payoffs to numerically calibrate their degree of aversion toward economic risk (Henrich & McElreath, 2002; Henrich & Smith, 2004). In neither society did measured risk preferences predict offers. Moreover, in both societies, subjects were risk preferring (formally, $\rho > 1$), rather than risk averse, a fact that casts further doubt on the risk aversion interpretation. We conclude that our offers are not explained by risk aversion *in the usual sense* (i.e., concave utility functions defined over gamble income, x^ρ with $\rho < 1$). Instead, high offers may reflect a desire to avoid rejections because of an aversion to social conflict, or a fear that a rejection is an awkward insult, rather than because of an aversion to variance in monetary outcomes (as in the economic model).

Alternatively, perhaps proposers are not sure how likely responders are to reject and offer more to be on the safe side. This tendency to behave cautiously in the face of unknown odds ("ambiguity" in economic language) is consistent with many other types of experimental data and economic phenomena (Camerer & Weber, 1992). In our settings, ambiguity-aversion toward rejection is plausible because the proposers do not see all the rejection frequencies that we observe. Whether ambiguity aversion can explain the high mean offers can be judged using the bootstrapping results shown in Figures 4a-b. That exercise produces 1000 different estimates of IMOs. Think of these as expressing the range of possible beliefs about rejections an uncertain proposer might entertain, and the optimal offers those wide-ranging beliefs imply. We can then ask: How pessimistic would proposers have to be to justify the mean offer as expected-income maximizing given pessimistic beliefs? A simple way to answer this question is to ask what fraction of the IMOs is above the mean offer. For most of the groups for which we can estimate IMOs at all, the results are striking: For the Achuar, Shona, Orma, Sangu herders, Machiguenga, and Mapuche, the mean offer is just slightly above the *most pessimistic* IMO among the 1000 simulated ones (which occurs when all the resampled offers are rejected). The mean offers/maximum IMO pairs are, respectively, 0.42/0.30, 0.44/0.40, 0.43/0.40, 0.41/0.33, 0.26/0.25, 0.335/0.33. It is as if subjects have a good guess about the highest offer that *could* be rejected, act as if that offer will be rejected *for sure*, and offer just above it to avoid rejection. Thus, while the gap between mean offers and IMO's visible in Figures 4a-b cannot be explained by risk aversion because of the concavity of the utility function for money, it can be explained as the result of pessimism about rejection frequencies and aversion to ambiguity.

For four groups (Aché, Tsimane, Kazakhs and Quichua) we could not estimate the IMO because there were no rejections. Nevertheless, as we discussed, it seems likely that substantially lower offers would have been accepted. Thus, offers in these groups cannot be explained by narrow self interest. Among the Au and Gnau, the IMO could not be established because responders from these groups did not preferentially accept higher offers, which is perhaps an even more striking violation of the selfishness axiom.

Additional evidence against the selfishness axiom comes from our three Dictator Games: the results here are more transparent than for the UG because the proposer is simply giving money away, anonymously, with no possibility of rejection. In each of the three groups in which the DG was played, offers deviate from the typical behavior of university students and from the predictions of self-regarding models. Mean offers among the Orma, Hadza, and Tsimane were 31, 20 and 32 percent, respectively, of the stake. These mean Dictator offers are 70, 60 and 86 percent of the corresponding mean UG offers for these groups. And, few or none of the subjects in these societies offered zero, while the modal offer among university students is typically zero (Camerer, 2003).[17]

Finally, the results from all six of our Public Goods Games also conflict with the Selfishness Axiom, with means ranging from 22% among the Machiguenga to 65% among the Aché—see Table 3. Even the Machiguenga data shows 62% of the sample violating the income-maximizing prediction of zero percent. Among the other groups, no group has more than 5% of the sample making contributions of zero. To our knowledge, this is never seen in one-shot PGGs among students, where a large percentage of players typically give zero.

Methodological variations between sites

Because our experiments were conducted at remote field sites with diverse, largely uneducated participants, we used some discretion in conducting the experiments to ensure comprehension and internal validity. The result was some methodological variation across sites.

For the UG, Table 4 documents the potentially important dimensions of variation in the administration of the experiments. These variations fall into eight categories. Beginning with column two, there were three different ways that the instructions used by different experimenters explained the allocation of the initial sum of money between the proposer and responder. In eight of our societies, the instructions stated that the money was allocated "to the pair," while in three societies the money was allocated "to the first person" (the proposer). Experimental economists have used both of these versions in their many UG experiments, and the results do not show any significant variation.[18]

Second, while most of our researchers stuck to entirely abstract explanations of the game and experimental context, using no explicit (and intentional) framing, two ethnographers did use some contextualization or framing in the games. To ensure comprehension among the Aché, Hill created an analogy between the UG and the process used by the Aché for apportioning the subcutaneous fat of game animals (Hill & Gurven, 2004). More indirectly, to attract Achuar and Quichua to the game, Patton called for a *Minga*, which, among these groups, are called to bring people together for cooperative work projects, like cutting a field for planting (Patton, 2004).

In a third kind of variation, five researchers read the instructions to a group first, and then brought the individuals into a gaming area to have their comprehension tested and make their decisions. Six others explained the games to individuals only after they entered the gaming area, and explained nothing to the group. Among the Machiguenga, Henrich (2000) used both methods and found no difference. Among university students this modification makes no difference.

Fourth, the difficulty of bringing everyone together at the same time led four researchers to conduct their experiments from house-to-house or one-by-one, sometimes spreading the games out over a few weeks. In seven other societies everyone was brought together in a single gaming area. Among the Machiguenga, both methods were used and no difference was found. Among students, this procedural variation does not impact the results (Henrich, 2000; Henrich & Smith, 2004).

Fifth, in all of our UG experiments, participants divided up sums of *cash*, except in Lamalera. There, to avoid the appearance of gambling, packs of cigarettes (which can be readily traded) were used as the medium of exchange instead of cash (Alvard, 2004).

Sixth, a few of our ethnographers, desiring to explore whether low offers would be rejected, fabricated offers for responders.[19]

Seventh, five groups were paid show-up fees for coming to the experiment (which subjects get regardless of what happens in the game), while seven others did not. U.S. research suggests that show-up fees do not have an important impact on UG play (Henrich & Smith, 2004; Henrich & Henrich, 2004).

Finally, one-on-one post-game interviews (to explore what people thought of the games, and why they did what they did, etc.) were done extensively in five societies, somewhat in four, and not at all in two groups. In one group, the Shona, Barr (2004) did focus groups.

Table 4. Summary of Methodological Variation across Field Sites

Site	Who the money was allocated to	Any explicit/intentional contextualization	Instructions to group first	Corralled players or house-by-house	Medium	Used any Deceptions	Show up fee	Post game interviews
Orma	the pair	No	Group	Corralled (no talking)	Cash	No	Yes	Some
Machiguenga	the pair	No	Both	Both	Cash	No	No	Yes
Mapuche	the pair	No	Individuals only	House-by-house	Cash	No	No	Yes
Au/Gnau	the first person	No	Individuals only	Corralled (talking)	Cash	No	Yes	No
Aché	the first person	Yes—related to meat sharing	Group	Corralled (talking)	Cash	Few sham low offers	Yes	Some
Tsimane'	the pair	No	Group	Corralled	Cash	No	Yes	Some
Lamalera	the pair	No	Group	Corralled (some talking)	Packs of Cigarettes	Sham low offers	No	No
Torguud Kazakhs	the first person	No	Individuals only	House-by-house	Cash	Sham low offers	No	Yes
Hadza	the pair	No	Individuals only	One-by-one (No corralling)	Cash	No	No	Yes
Shona	ambiguous	No	Individuals only	Corralled (No talking)	Cash	No	No	Group debriefs
Achuar Quichua	the pair	Yes—people invited to a "Minga"	Group	Corralled (No talking)	Cash	No	Yes	Some
Sangu	the pair	No	Individuals only	Both	Cash	No	No	Yes

Three reasons suggest that these methodological variations cannot account for the broad patterns of variation we observed. First, there is no reliable correspondence between methodological variations across groups in the UG and their game behavior (compare Tables 2 and 4). Second, as noted, many of these variations do not produce substantial variations in the populations where they have been tested.[20] Third, in several cases in which the identical protocols and experimenters were used in different places, the results still show substantial variation. The following subsets faced the *identical* experimenters and protocols and still showed substantial variation: (1) Machiguenga, UCLA students (a student control, see Henrich, 2000) and the Mapuche (Henrich: these three yielded UG mean offers of 26, 48, and 34 percent), (2) the Quichua and Achuar (Patton: UG mean offers of 25 and 43 percent). The same can be said of the PGG data, where identical protocols and experiments were used for the Machiguenga vs. Michigan, and the Huinca vs. Mapuche comparison. Moreover, within our linguistic groups, individual researchers found substantial variation between communities (Tsimane, Sangu, Shona and Hadza)—more on this below. By the same token, however, the same experimenters and protocols did not always find between-group variation, as these comparisons attest; (1) Kazaks and Mongols (Gil-White), and (2) the Au and Gnau (Tracer).

It is also important to realize that UG results from industrialized societies are generally quite robust against a wide range of procedural variations (that's why we selected it for the project!).[21] While many experimentalists have highlighted 'significant differences' in framing effects for the UG, the size of these differences is almost always small compared to the kinds of differences we found cross-culturally (Camerer, 2003:

Chapter 2). Thus, 'significant effects' should not be confused with big effects (and, one should also consider that treatments that result in non-significant differences will rarely see the light of day). The largest of these effects (among university students) involves substantial manipulations, like including a pre-game trivia contest that determines who gets to be the proposer. Under these conditions, proposers offer less, and responders accept less (Hoffman, McCabe, Shachat, & Smith, 1994). Certain contextualizations (e.g., a monopoly seller choosing a price) have a modest effect on offers, shifting the mean by about 10% of the pie (Camerer, 2003: Chapter 2; Hoffman et al., 1994). Other seemingly important variations actually have little effect on offers (Larrick & Blount, 1997). Playing repeatedly (with feedback about one's own results) or increasing stakes by up to a factor of 25 changes offers by only 1-2%. In contrast, moving the identical protocol from the Machiguenga to UCLA increases offers by 85%.

It is important to realize that the few variations in UG instructions or procedures that have shown a substantial impact on results were *deliberately* designed by researchers because they suspected that such variations might cause a big effect. In contrast, our researchers tried to avoid any modifications that might have an effect, and our variations were typically ad hoc procedures created by field researchers in adapting to the field situation, or inadvertent nuances due to translation. Such variations, for example, do not result in accidentally slipping a trivia contest (that determines who gets to decide the monetary division) into the instructions.

A final methodological concern in interpreting the cross-cultural results comes from possible experimenter bias. The relationships between our experimenters and the participants are typically much closer, more personal, and longer lasting than in university-based experiments. Consequently, it is possible that ethnographers may bias the results of our experiments in ways different from that found in standard situations. Two pieces of data argue against this interpretation. First, Henrich (2000) attempted to control for some of this effect by replicating the Machiguenga UG protocol with UCLA graduate students. In this control, Henrich and his subjects knew one another, had interacted in the past, and would interact again in the future. His results were quite similar to typical UG results in high stakes games among adults in the U.S., and substantially different from the Machiguenga. This is certainly not a complete control for experimenter bias, but it does confront some elements of the bias. Second, to test for experimenter bias across our samples, we examined the relationship between the time each experimenter had spent in the field prior to administering the games and the mean UG of each group, but found no consistent pattern in the data. Finally, since most people would predict that having some longer-term relationship with the experimenter would bias offers towards generosity, and most of our variation is more selfish than university student results, it is difficult to argue that such a bias is driving the results. Nonetheless, we cannot entirely exclude the possibility that some of the observed between-group differences result from differences among the experimenters and the details of how the experiments were implemented.

(4) Explaining Group Differences in Behavior

To examine the variation between groups, we first examined whether any attributes of individuals were statistically associated with proposer offers across our sample. Among the measured individual characteristics that we thought might explain offers were the proposer's sex, age, level of formal education, and their wealth relative to others in their group.[22] In pooled regressions across all offers none of these individual-level variables predicted offers once we allowed for group level differences in offers (by introducing group dummy variables). Since the group dummy variables account for approximately 12% of the variance in individual offers, we conclude that group differences are important. However, for the moment, we remain agnostic about the role of individual differences. Our pooled regression tested for common effects of these variables across all the groups and hence does not exclude the possibility that the individual differences we have measured may predict behaviors in different ways from group to group. We return to this below.

In proposing this project, we hypothesized that differences in economic organization and independence, social organization (complexity), and market integration may influence cultural transmission and create between-group differences in notions of fairness and punishment.[23] To test these initial hypotheses, we rank ordered our societies along five dimensions. First, *Payoffs to Cooperation* (PC)—to what degree does economic life depend on cooperation with non-immediate kin? In a sense, PC measures the presence of extra-familial cooperative institutions. Groups like the Machiguenga and Tsimane ranked the lowest because they are almost entirely economically independent at the family level. In contrast, the economy of the whale hunters on Lamalera depends on the cooperation of large groups of non-kin. Second, *Market Integration* (MI) —do people engage frequently in market exchange? Hadza foragers were ranked low because their life would change little if markets suddenly disappeared. Others, like the Orma, were ranked higher because they frequently buy and sell livestock and work for wages. Third, *Anonymity* (AN)—how important are anonymous roles and transactions? While many

Achuar of the Ecuadorian Amazon never interact with strangers, the Shona of Zimbabwe frequently interact with people they do not know and may never see again. Fourth, *Privacy*—how well can people keep their activities secret from others? In groups like the Au, Gnau and Hadza, who live in small villages or bands and eat in public, it's nearly impossible to keep secrets and quite difficult to hide anything of value. Among the Hadza, simply having pants increases privacy because they have pockets. In contrast, Mapuche farmers live in widely scattered houses and maintain strict rules about approaching another's house without permission, so privacy is substantial. Fifth, *Sociopolitical Complexity* (SC)—how much decision-making occurs above the level of the household? Because of the importance in the anthropological literature of the classifications of societies by their political complexity (Johnson & Earle, 2000), we ranked our societies from family level through chiefdoms and states. Finally, *Settlement Size* (SS)—the size of local settlements, which ranged from less than 100 among the Hadza to more than 1000 on Lamalera.

Before beginning the data analysis we ranked the groups along these dimensions using the following procedures. First, during a meeting of the research team, we had a lengthy discussion of the underlying attributes that each dimension was designed to capture. Then the field researchers lined up and sorted themselves by repeatedly comparing the group they studied with those of their two neighbors in line, switching places as necessary, and repeating the process until no one needed to move. The subjective nature of the resulting ordinal measures is evident.[24] Second, our complexity rankings were generated by both Henrich (who was not blind to our experimental results) and Allen Johnson, an outside expert on societal complexity, who was blind to the results. Henrich's and Johnson's rankings correlated 0.9, and explain about the same amount of variation in mean UG offers.

We have no way of knowing the direction of causality between the measures of social structure and offers. An ideal way to disentangle causality is to have an exogenous variation in structural conditions and correlate it with offers (what econometricians call an "instrumental variable"). The time course of history in these societies does not permit such an inference.

Table 5 Correlation matrix for our group-level variables

	PC	AN	MI	PR	SS
Social Complexity (SC)	.242	**.778**	**.913**	.374	**.670**
Payoffs to Cooperation (PC)	–	–.063	.039	–.320	.165
Anonymity (AN)		–	**.934**	**.743**	**.664**
Market Integration (MI)			–	**.644**	**.731**
Privacy (PR)				–	.328
Settlement Size (SS)					–

As seen in Table 5, four of these indices, market integration, anonymity, social complexity, and settlement size, are highly correlated across groups suggesting that they may all result from the same underlying causal process. The correlation of each of these variables with the potential payoffs to cooperation is very small suggesting that this ranking measures a second set of causal factors. This is not surprising. An increase in social scale is associated with a shift to market based economy, and an increase in anonymity. Within small scale societies with similar levels of social complexity, there is a wide range of economic systems with varying levels of cooperation. To capture the causal effects of this nexus of variables, we created a new index of "aggregate market integration" (AMI) by averaging the ranks of MI, SS, and SC (We did not include AN because it is so similar to MI, and including it has only a slight effect).

We estimated ordinary least squares regression equations for explaining group mean UG offers using the payoffs to cooperation (PC) and aggregate market integration (AMI). Both of their normalized regression coefficients are highly significant and indicate that a standard deviation difference in either variable is associated with roughly half a standard deviation difference in the group mean offers (Table 6, Figure 5). Together these two variables account for 47% (adjusted R^2) of the variance among societies in mean UG offers. The magnitude of these coefficients, and their significance, is robust to three different checks on the analysis.[25]

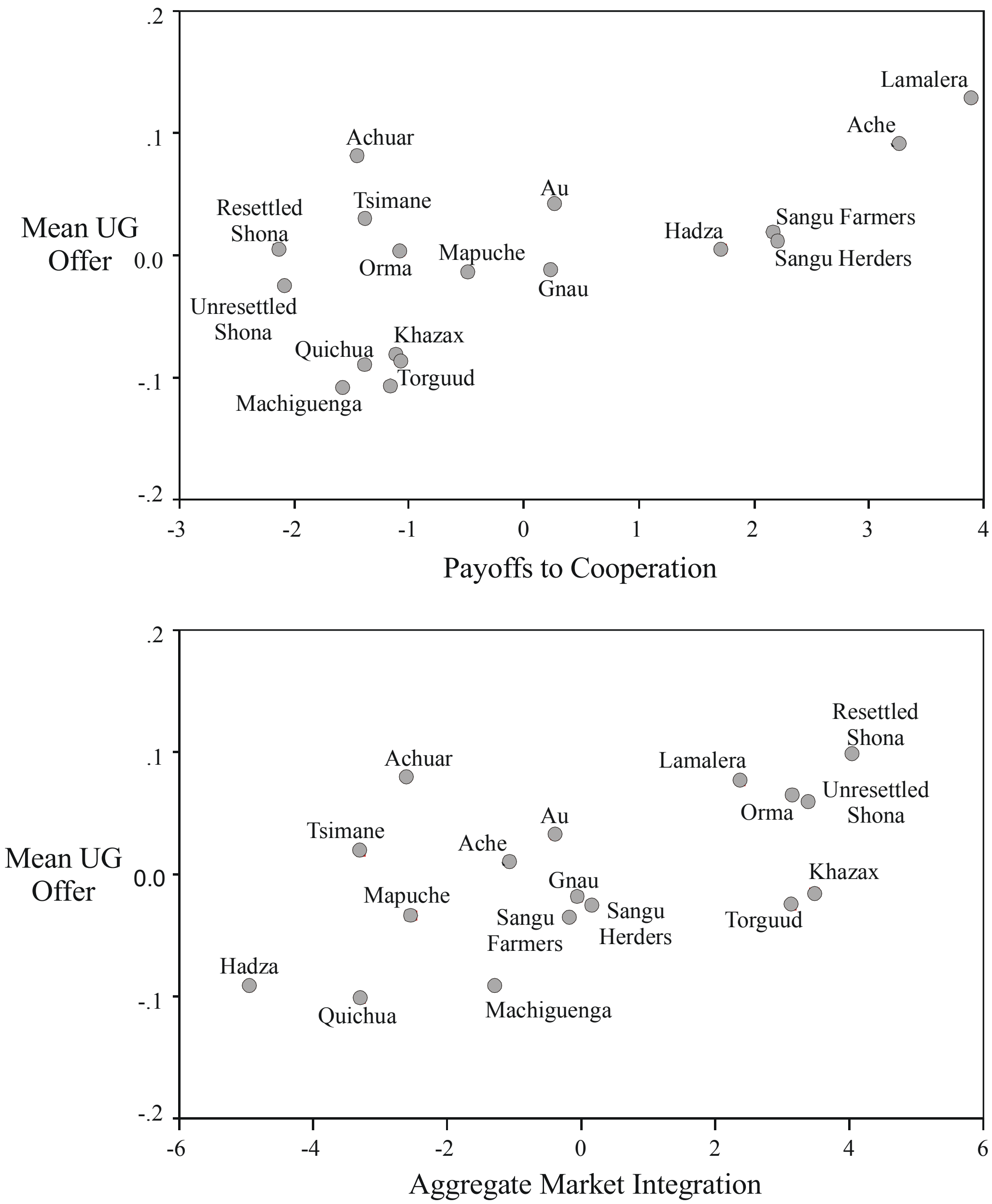

Figure 5. Plots of mean UG offers as a function of the PC and AMI indices. The vertical and horizontal axes measure units of standard deviation in the sample. Because AMI and PC are almost uncorrelated (r =.04), these bivariate plots give a good picture of their effects.

Table 6. Regression coefficients and statistics

	Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	.261	.036	--	7.323	.000
CP	.021	.007	0.528	2.922	.011
AMI	.012	.005	0.448	2.479	.027

All regressions using PC and one of the other predictors (AN, MI, SC, and SS) yielded a significant positive, coefficient for PC and a positive, nearly significant coefficient for the other variable. If we use the IMO (Income Maximizing Offer) as a predictor of the UG offers along with PC and AMI, its coefficient is smaller (in magnitude), negative and insignificant, while the coefficients of PC and AMI remain large and close to significance at conventional levels (even though for IMO $n = 9$), suggesting that the effects of economic structure and cultural differences captured by PC and AMI do not substantially influence offers through the IMO.

The same two variables (PC and AMI) also predict the group average IMO; the effect sizes are large (normalized regression coefficients about one half) but very imprecisely estimated (significant only at the 20% level). Taken at face value, these estimates suggest that subjects' expectations about the likelihood that low offers will be rejected covaries with both the benefits of cooperation and aggregate market integration.

Our analysis of the individual level responder data across all groups reveals some of the same basic patterns observed in the proposer data. The age, sex, and relative wealth of a responder does not affect an individual’s likelihood of rejecting an offer. What does matter is the proportion of the stake offered and the responders’ ethnolinguistic group.

(5) Explaining Individual Differences within Groups

In contrast to the power of our group level measures in statistically explaining between group differences in experimental behaviors, our individual level variables explain little of the variation within or across groups. With a few group-specific exceptions, nothing we measured about individuals other than their group membership (society, village, camp, or other subgroup membership) predicted experimental behavior. Here we summarize our findings concerning individual attributes and experimental play in *within group* analyses. Sex, wealth, and age do not generally account for any significant portion of the variance in game play. However, in the UG, sex was marginally significant among the Tsimane, where males offer 10% more than females (Gurven, 2004). Among the Hadza, women's UG offers strongly increased with camp population size, but camp size was not important to men's offers. Conversely, in the DG, it was the offers of Hadza men that increased with camp size (Marlowe 2002). As in the UG, Public Goods Game data from five societies also reveal no significant effects of sex, except among Aché men who contribute a bit more than women (Hill & Gurven, 2004). Similarly, wealth—in any form (e.g. cash, cows, land)—fails to predict game behavior. In several circumstances, multiple measures of wealth (e.g. animal wealth, cash, and land-wealth) were gathered and analyzed, as well as an aggregate measure. In these within-group analyses, wealth arose as significant only once in 12 different data sets, including both UG and PGG games. The exception comes from an all-male Public Goods Game among the Orma. Controlling for age, education, income, and residence pattern (sedentary vs. nomadic), wealth was the only significant predictor of contributions in a multivariate linear regression, with a standard deviation difference in wealth predicting well over half a standard deviation increase in contributions (Ensminger, 2004)—we make sense of this finding below.

Several researchers also analyzed measures of formal education. Analyzing UG data from the Sangu, Orma, Mapuche, Au and Gnau, we find that the extent of schooling does not account for any significant portion of the variation in offers in either bivariate analyses or multivariate regressions that controlled for sex, age and wealth. Among the Tsimane, the extent of formal education emerges as marginally significant in a multivariate regression involving age, village, sex, Spanish-speaking ability, trips to the nearest market town, and wage labor participation. Less educated Tsimane offer *more* in the UG game. However, we find no effect of formal education on PGG play among the Tsimane. Thus while schooling effects may exist, they are neither particularly strong nor consistent across games or societies.

Although our group level measure of market integration has impressive statistical power, individual level measures of market exposure do not explain any significant proportion of the variation within groups. To assess market integration, some researchers gathered data on individuals' participation in wage labor, their reliance on cash cropping, and their competence in the national language. Wage labor participation shows no significant relation to UG offers in six groups—the Tsimane, Aché, Gnau, Au, Machiguenga and Mapuche. PGG data from the Orma, Aché, Machiguenga and Tsimane also show that wage labor does not influence game play. The only clear exception to the wage labor pattern occurs in the Orma UG data, where individuals who participate in wage labor (to any degree) make significantly higher offers than those who do not (Ensminger, 2004).

In societies based on agriculture, another measure of market integration is the amount of land an individual (or household) devotes to cash cropping, as opposed to subsistence cropping. We obtained cash cropping data from three societies. Among the Machiguenga, land (in hectares or as a proportion of total land) devoted to cash cropping is positively correlated with UG offers; its normalized partial regression coefficient when age, sex and wage labor are controlled remains substantial, though its significance level is marginal (Henrich & Smith, 2004). Neither total cash-cropping land nor the proportion of land devoted to cash cropping is significantly related to UG offers for the Au and Gnau. However, among the Au (but not the Gnau) multivariate regressions show that land devoted to *subsistence* cropping positively predicts UG offers, controlling for sex, age, cash cropping land and wage labor (Tracer, 2003, 2004).

In many places, an individual's degree of competence in the national language may also represent a measure of market integration, or at least of market exposure. We only have language data from the Tsimane. While it is significant in bivariate analyses, multivariate regressions that control for village membership, sex, age, visits to San Borja, years of formal education, and participation in wage labor, show no relationship between Spanish-speaking ability and UG offers. Using the same controls, competence in the national language also fails to predict PGG contributions (Gurven, 2004).

As is the case for all of our individual level data, except for age and sex, these measures capture individual behaviors that may well be endogenous with respect to the beliefs or preferences our experiments measure. Because it is possible that these measures are the consequence, rather than the cause of individual behavioral differences, we also sought to use geographical measures of proximity to market opportunities as exogenous instruments for measuring market exposure in three groups: Tsimane, Au and Gnau. None of these were significant predictors of proposer behavior.

It is possible, of course, that the unexplained within-group variance in experimental behaviors reflects a lack of comprehension of the game or errors in experimental play that are unrelated to measures like age, wealth or wage labor participation. Overall, we have little reason to suspect that game comprehension significantly influenced the results (although see Gil-White, 2004). In most cases experimenters tested subjects for game comprehension before the experiments were implemented, and excluded those who had difficulty grasping the game. In several studies, experimenters used post-game interviews to probe for possible misunderstandings and faulty assumptions. Among the Mapuche, the players who passed the basic tests were ranked according to how well they understood the strategic nature of the game, and how well they were able to do the monetary calculations involved. Neither measure predicts game behavior or deviation from mean game behavior. Similarly, among the Hadza (F. Marlowe, 2004), players were scored according to the number of practice examples it took for them to learn the game. Among Hadza males this measure is unrelated to both UG proposer and responder behavior, but for females comprehension is positively and significantly correlated with offer size. We do not know if the covariation of comprehension and experimental behavior among Hadza women represents the effect of comprehension per se, or results from the association of comprehension with other correlates of game play for women, such as camp size (a strong predictor of Hadza women's offers). Finally, as noted above, education—which might be thought to correlate with degree of game comprehension—did not predict game behavior.

Given that we sought individual level statistical associations for a number of variables in 15 societies and found just a handful of estimates suggesting substantial effects, we conclude that, other than group membership, the individual-level facts we have collected about our subjects do not consistently predict how individuals will behave. This does not mean that within group variation in subjects' behavior cannot be explained; rather it suggests that the explanation may be group-specific and/or that we may not have collected the appropriate information.

(6) Local Group Effects

Our analysis suggests that group-effects may be important and this opens the question of how to define a group. In the above analyses, ethnolinguistic markers were used to define group membership, but non-ethnolinguistic regional groupings, or smaller local clusters (e.g., villages) may be more appropriate. Our data allow some comparisons. Such small-scale tests permit us to control for a number of variables, including climate, language, regional/national economy, local buying power of the game stakes, and local history. In the Bolivian Amazon, the effects of market integration and local-groups were examined by performing the UG and PGG in five different villages at different distances from the market town of San Borja, the only source of commercial goods, medicines, and wage labor opportunities. Like the Machiguenga, the Tsimane live in small communities scattered along a major riverine drainage system. Under these circumstances, physical distance (in travel time along the river) from San Borja provides a proxy measure for the extent of market contact of different Tsimane communities. As noted, the results indicate that a community's distance from San Borja is unrelated to UG or PGG behavior. Interestingly, however, the best predictor for UG offers and PGG contributions is what community one is from, *independent* of the community's distance from San Borja and population size. So, where a Tsimane lives matters, but differences in both individual-level measures of market integration and community-level market variables apparently do not. Among the Tsimane, the relevant group for predicting UG and PGG behavior appears to be smaller than the ethnolinguistic group.

Like Tsimane, we found a number of cases in which group membership effects were strong even in the absence of geographical isolation, suggesting that the processes that generate and maintain behavioral differences among groups can maintain differences between frequently interacting, and even intermarrying, groups. In Chile, Mapuche farmers and non-Mapuche Chilean townspeople, locally called Huinca, have lived side-by-side, intermarried and interacted for at least 100 years. Yet, the Mapuche and the Huinca behave quite differently in a single-shot PGG game. The Mapuche contributed a mean of 33% to the pot, while the Huinca offered an average of 58%. In Ecuador the Achuar and Quichua of Conambo, who interact and intermarry frequently, play the UG quite differently: Achuar proposers offered a mean of 43% while Quichua proposers offered only 25%. This difference is especially notable as Quichua and Achuar subjects were randomly paired, so the proposers from the two groups faced the same probability of rejection. In Tanzania, Hadza from the biggest camp (which was three times larger than the next largest camp) played the UG much more like university students than like Hadza from the four smaller camps, despite the fact that camps are ephemeral social units and camp membership is quite fluid. For the Hadza, camp population size turns out to be the best predictor of UG offers—the larger the camp, the higher the mean UG offer. Finally, although Sangu herders and farmers make similar UG offers, farmers reject offers more frequently than herders. Yet, Sangu often change from herder to farmer and back again over the course of one lifetime.

In contrast, local groups in some locations showed no or little between-group variation. In Mongolia, Torguud Mongols and Kazakhs are separated by deep cultural and historical differences, yet they play the UG similarly. In Papua New Guinea, the Au and Gnau, who speak mutually unintelligible languages and show differing degrees of market incorporation, played the UG in the same unusual manner (making and rejecting offers over 50%). Similar comparisons in Zimbabwe between resettled and unresettled Shona reveals only slight differences.

In general, the micro-level variation we observed contrasts with the UG results from the U.S. and Europe in which university students, who speak different languages and live thousands of miles apart, behave quite similarly. Of course, it is possible that variation exists within contemporary societies, but this variation is not represented in university populations (Cummings & Ferraro, 2002). Nevertheless, recent UG experiments with subjects outside of universities in the U.S. have failed to uncover behavioral patterns in the UG much different from those observed among university students (Carpenter et al., 2002; Henrich & Henrich, 2004).

(7) Experimental Behavior and Everyday Life

The fact that group level measures of economic and social structure statistically explain much of the between group variance in experimental play suggests that there may be a relationship between game behavior and patterns of daily life in these places. In several cases the parallels are striking, and in some cases our subjects readily discerned the similarity, and were able to articulate it. The Orma, for example, immediately recognized that the PGG was similar to the *harambee,* a locally-initiated contribution that Orma households make when their community decides to pursue a public good, such as constructing a road or school. They dubbed the experiment "the *harambee* game" and contributed generously (mean 58% with 25% full contributors).

Recall that among the Au and Gnau of Papua New Guinea many proposers offered more than half the pie, and many of these offers were rejected. The rejection of seemingly generous offers, of more than half, may have a parallel in the culture of status-seeking through gift giving found in Au and Gnau villages, and throughout Melanesia. In these societies, accepting gifts, even unsolicited ones, implies a strong obligation to reciprocate at some future time. Unrepaid debts accumulate, and place the receiver in a subordinate status. Further, the giver may demand repayment at times, or in forms (political alliances), not to the receiver's liking—but the receiver is still strongly obliged to respond. As a consequence, excessively large gifts, especially unsolicited ones, will frequently be refused. Together, this suggests that as a result of growing up in such societies, individuals may have acquired values, preferences or expectations that explain both high offers and the rejection of high offers in a one-shot game. Interestingly, it may turn out that what is unique here is not the rejection of high offers (ethnographically, many societies disdain excess generosity), but the willingness to make offers over 50%.

Among the whale hunting peoples on the island of Lamalera (Indonesia), 63% of the proposers in the Ultimatum Game divided the pie equally, and most of those who did not, offered more than half (the mean offer was 58% of the pie). In real life, when a Lamalera whaling crew returns with a large catch, a designated person meticulously divides the prey into pre-designated parts allocated to the harpooner, crewmembers, and others participating in the hunt, as well as the sail maker, members of the hunters' corporate group, and other community members (who make no direct contribution to the hunt). Because the size of the pie in the Lamalera experiments was the equivalent of ten days wages, making an experimental offer in the UG may have seemed similar to dividing a whale.

Similarly, in Paraguay the Aché regularly share meat. During this sharing, the hunters responsible for the meat forgo their share, while the prey is distributed equally among all other households. There is no consistent relationship between the amount a hunter brings back and the amount his family receives (Kaplan & Hill, 1985). Successful hunters often leave their prey outside the camp to be discovered by others, carefully avoiding any hint of boastfulness. When asked to divide the UG pie, Aché proposers may have perceived themselves as dividing the game (meat) they or a male member of their family had acquired, thereby leading 79% of the Aché proposers to offer either half or 40%, and 16% to offer more than 50%, with no rejected offers.

By contrast, the low offers and high rejection rates of the Hadza, another group of small-scale foragers, are not surprising in light of the numerous ethnographic descriptions (F. Marlowe, 2004; Woodburn, 1968). Although the Hadza extensively share meat (and other foods to a lesser degree), they do not do so without complaint, and many look for opportunities to avoid sharing. Hunters sometimes wait on the outskirts of camp until nightfall so they can sneak meat into their shelter (F. W. Marlowe, 2004). The Hadza share because they fear the social consequences that would result from not sharing. Cooperation and sharing are enforced by a fear of punishment that comes in the form of informal social sanctions, gossip, and ostracism (Blurton Jones, 1984, 1987). Many

Hadza proposers tried to avoid sharing, and several of them were punished by rejection. Thus, we find two foraging peoples—the Aché and the Hadza—at opposite ends of the UG spectrum in both offers and rejections, with each seeming to reflect their differing patterns of everyday life.

Similarly, both 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. The Machiguenga, for example, did not even have personal names until recently—presumably because there was little reason to refer to people outside of one's kin circle (Johnson, 2003). Consequently, it's not very surprising that in an anonymous interaction both groups made low UG offers. Given the Tsimane UG offers vary across villages, it would be interesting to know if these differences reflect village-level differences in real prosocial behavior.

While methodological discussions commonly address the correspondence of experimental regularities to behavior in naturally-occurring economic interactions (Camerer, 1996; Loewenstein, 1999), our concern here is more modest: to explore the possibility of a connection between patterns of behavior in the experiments and those in the daily lives of our subjects. In many societies it appears that there may be such a connection, and that sometimes our subjects were able to verbalize those parallels.

(8) Discussion: Theoretical Implications

Understanding the patterns in our results calls for incorporating proximate-level decision-making models from behavioral economics, which have increasingly drawn insights on human motivation and reasoning from psychology and neuroscience

(Camerer, 2003; de Quervain et al., 2004; Sanfey et al., 2003), under the ultimate-level evolutionary umbrella created by culture-gene coevolutionary theory (Baldwin, 1896; Boyd & Richerson, 1985; Campbell, 1965; Cavalli-Sforza & Feldman, 1981; Durham, 1991; Pulliam & Dunford, 1980). Coevolutionary theory treats genes and culture as intertwined informational systems subject to dual evolutionary forces. In our species, cultural capacities are best understood as sophisticated social learning mechanisms (Tomasello, Carpenter, Call, Behne, & Moll, in press) for acquiring—at low cost—locally adaptive behaviors or decision information. Because these forms of social learning create cumulative evolutionary products over generations (e.g., technologies), multiple stable equilibria in social interactions (e.g., institutional forms), and operate on much shorter time scales than genetic evolution (Boyd & Richerson, 1996; Gintis, 2003a; Tomasello, 1999), cultural evolution and its products have undoubtedly influenced the human genotype (Bowles & Gintis, 2004). This theoretical avenue predicts that humans should be equipped with learning mechanisms designed to accurately and efficiently acquire the motivations for the local set of culturally-evolved social equilibria (institutions).

Behavioral game theory—the subdiscipline from which our methods derive—is rooted in the notion that individuals will select among alternatives by weighing how well the possible outcomes of each option meet their goals and desires. Theoretically, this is operationalized by assuming agents maximize a *preference function* subject to informational and material *constraints*. Behavioral game theory shows that by varying the constraints and the rewards, as assessed by the agent's preference function, as we do in such games as the UG and PGG (Charness & Rabin, 2000; Fehr & Schmidt, 1998), we can determine the arguments of the agent's preference function and how the agent trades off amongst desired rewards. We call this the preferences, beliefs and constraints approach.

It is often thought that the preferences, beliefs and constraints approach presumes that individuals are self-regarding, and/or that they have very high levels of reasoning or omniscience. However, while this has often been true of many models, these assumptions are certainly not necessary. Indeed our research (along with much other work) shows that such considerations as fairness, sympathy, and equity are critical for understanding the preference functions of many humans, and can be effectively integrated with such things as pleasure, security, and fitness to produce a more complete understanding of human behavior. Similarly, these models do not necessarily presume anything in the way of reasoning ability, beyond that required to understand and perform in everyday social contexts.

The relationship between culture-gene coevolutionary theory and the preferences, beliefs and constraints approach is straightforward, although rarely illuminated. As background, evolutionary game theory has shown that social interactions among populations of individuals with adaptive learning mechanisms often produce multiple stable social equilibria (Fudenberg & Levine, 1998; Gintis, 2000; Weibull, 1995; Young, 1998). Different human ancestral groups through time, as they spread across the globe and adapted their behavioral repertoire to every major habitat from the malarial swamps of New Guinea to the frozen tundra of the Siberian Arctic, would have culturally evolved different social equilibria (forms of social organizations and institutions).[26] As a consequence, ancestral humans would have needed to adapt themselves ontogenetically to the vast range of potential social equilibria that one might encounter upon entering the world. The result of dealing with this adaptive problem, we argue, is that humans are endowed with cultural learning capacities that allow us to acquire the beliefs and preferences appropriate for the local social environment—i.e., human preferences are *programmable*, and are often internalized just as are aspects of our culinary and sexual preferences. The *preferences* become part of the preference function that is maximized in preferences, beliefs and constraint models. Norms such as 'treat strangers equitably' thus become valued goals in themselves, and not simply because they lead to the attainment of other valued goals.

The theory sketched above has two immediate empirical entailments. First, people should rely on cultural learning to acquire significant components of their social behavior. If they don't, the theory cannot even get off the ground. Second, as a consequence of these adaptive learning processes, societies with different historical trajectories are likely to arrive at different social equilibria. As such, people from different societies will tend to express different preferences and beliefs: one should be able to measure between-group variation. With regard to this second entailment, we submit the above results from our cross-cultural project.

For the first entailment, there is ample evidence from psychology and sociology that humans acquire much of their social behavior through cultural learning. Psychologists have amassed evidence showing that children spontaneously (without incentives) acquire social behavior by observing and imitating others (Bandura, 1977; Rosenthal & Zimmerman, 1978). More to point, studies of prosociality in children show that children readily imitate models demonstrating either costly altruism or selfishness

(Bryan, 1971; Bryan & Walbek, 1970; Grusec, 1971; Presbie & Coiteux, 1971). Additional work demonstrates that (1) this effect is not ephemeral, and can be seen in re-tests months later (Rice & Grusec, 1975; Rushton, 1975), (2) the effect is increased somewhat if values are strongly voiced along with actions (Grusec, Saas-Kortsaak, & Simutis, 1978; Rice & Grusec, 1975; Rushton, 1975), (3) sometimes these imitation patterns are generalized to other quite different contexts (Elliot & Vasta, 1970; Midlarsky & Bryan, 1972), and (4) children used learned standards of altruism to judge and punish others (Mischel & Liebert, 1966). Some of the details of how norms get internalized has been studied in socialization theory (Grusec & Kuczynski, 1997; Parsons, 1967).

Integrated with these basic cultural processes, the preferences and beliefs of new members are influenced by the economic and social institutions that structure the tasks people perform to make a living and to remain in good standing in their communities. Indeed, evidence from experiments, industrial sociology, and ethnography suggest that commonly performed tasks affect the basic values incorporated in the individual's preference function, and hence will be expressed far beyond the limits of the workplace or the specific institutional structure responsible for their social prominence. In experimental work, Sherif (1937) and others have shown that the performance of cooperative tasks (in which success depends on the efforts of many and the rewards are shared) induces positive sentiments toward those with whom one cooperates. Competitive tasks produce the opposite effect. Sociological and ethnographic studies show that the degree of autonomy one exercises in making a living, for example, is strongly associated with child rearing values in industrial (Kohn, 1990) and small-scale (Barry, Child, & Bacon, 1959) societies. That these values are widely internalized and expressed is exemplified by the fact that group-level average UG offers, and PGG contributions, are highly correlated across the societies in which both games were played ($r = 0.79$, $p = 0.06$, $n = 6$).

Consistent with this view is evidence from UG, DG, and PGG experiments among children and adults in the U.S. showing that preferences related to altruism, conditional cooperation and equity are acquired slowly over the first two decades of life (second graders are pretty selfish), and subsequently change little after this (Harbaugh & Krause, 2000; Harbaugh et al., 2002; Henrich, 2003).

Because of the nature of our adaptive learning processes, individuals in experiments bring the preferences and beliefs that they have acquired in the real world into the decision-making situation. The social relations of daily life may lead individuals to generalize about how others are likely to act in novel situations. Thus, for instance, if there is a high level of cooperation in work and/or community, people may expect others to behave in a similarly cooperative manner in novel situations, such as those provided by experimental games. If people prefer to cooperate when others cooperate (as shown by experimental data Fehr & Gächter, 2000; 2002 and in cross-cultural data ; Henrich & Smith, 2004), and if they have reason to believe others will cooperate, they themselves will likely cooperate, thus leading to a high level of cooperation in the experimental situation. If subjects believe others will not cooperate, and even if they prefer to cooperate as long as others do so as well, a low level of cooperation will likely result. For example, participants in a market-oriented society may develop distinct cognitive capacities and habits. Moreover, extensive market interactions may accustom individuals to the idea that strangers can be trusted (i.e., expected to cooperate). This idea is consistent with the fact that UG offers and the degree of market integration are strongly correlated across our groups.

Demonstrating the effect of contextual interpretation on beliefs and expectations, experiments with students in industrialized societies have shown that contextual cues can significantly change contributions in social dilemmas. This dramatizes the importance of expectations in strategic cooperative behavior. For example, Ross and Ward (1996) and Pillutla and Chen (1999) used two versions of a public goods game—one construed as a joint investment or "Wall Street game," and the other as a contribution to a social event or "community game." Players contributed significantly less to the investment than to the social event, holding their payoff structures constant (also see Hayashi, Ostrom, Walker, & Yamagishi, 1999).[27]

For some cues, culture and context interact. Cues that create an effect in one place do not create the same effect elsewhere. For example, Kachelmeier & Shehata (1997), in a public goods experiment comparing Canadian, mainland Chinese and Hong Kong students, showed that low anonymity conditions led Chinese students, especially Mainlanders, to behave very cooperatively while having no effect on Canadians. Similarly, Hayashi et. al. (1999) showed that certain framing effects strongly influence cooperation rates among Japanese students, but not among Americans.

The details of how daily life enters the experimental situation to influence behavior remain unclear. Two non-exclusive possibilities deserve note. It may be that different social, cultural and physical environments foster the development of differing *generalized behavioral dispositions* (equity, altruism, etc.) that are applicable across many domains, as might be the case using the above reasoning concerning task performance or investment in reputation building. For example, the Lamalera may be generally more 'altruistic' or 'fair-minded' than Machiguenga or Quichua. In our experimental situations, such dispositions could account for the statistical relationships between group characteristics and experimental outcomes. Alternatively, the abstract game structures—which are standard in such experiments—may cue one or more highly context specific behavioral rules (or sets of preferences), as is suggested by the situational framing examples above. In these situations, subjects in some places were first identifying the kind of situation they were in, seeking analogues in their daily life, and then acting appropriately. In this case, individual differences result from the differing ways that individuals frame a given situation, not from generalized dispositional differences. Given what is known about how generalized values develop, it's plausible that both are going on to differing degrees in different societies.

One of our cases allows a distinction between the two. Recall that the Orma made a connection between the public goods game and their local practice, the *harambee.* The Orma believe that wealthier households should make larger contributions to the *harambee* than poorer households. The Orma did not perceive a similar connection between the *harambee* and the UG. Multivariate regressions involving wealth, age, education and income indicate that wealth is the only significant predictor of PGG contributions among Orma individuals. The more wealth a person has the more they contribute to the common pool, just like in the real *harambee*. Wealth, however, is not a significant predictor of UG offers in either multivariate or bivariate analyses. The importance of wealth for PGG games, but not for UG, is consistent with predictions from the context specific approach, assuming that the resemblance of the public goods game to the familiar *harambee* cues appropriate behavior in that game but does not generalize to the uncued.

Combining a preferences, beliefs and constraints approach with culture-gene coevolutionary theory produces a framework that endogenizes both the cultural and genetic aspects of human preferences and beliefs, while at the same time retaining analytically tractable models that permit quantitative predictions of behavior (Camerer, 2003; Fehr & Schmidt, 1998; Fischbacher, Fong, & Fehr, 2002). Coevolutionary approaches provide a firm theoretical foundation for studying the psychological mechanisms that permit us to rapidly and accurately acquire the locally adaptive preferences, norms and beliefs (Gintis, 2003a, 2003b; Henrich & Gil-White, 2001; Richerson & Boyd, 2000). Cultural evolutionary models allow us to explore the conditions and processes that generate different preferences and beliefs that form the foundation for the diversity of human institutions and social norms (Boyd, Gintis, Bowles, & Richerson, 2003; Henrich & Boyd, 2001; McElreath, Boyd, & Richerson, 2003). Each of these evolutionary processes helps us to understand where the preferences and beliefs—the critical ingredients of the decision-making models—come from, and how they have evolved over human history, on both shorter and longer time scales (Bowles, 1998).[28]

References Cited

Alvard, M. (2004). The Ultimatum Game, Fairness, and Cooperation among Big Game Hunters. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Experiments and Ethnography in 15 small-scale societies* (pp. 413-435). London: Oxford University Press.

Andreoni, J. (1988). Why Free Ride - Strategies and Learning in Public Goods Games. *Journal of Public Economics, 37*, 291-304.

Andreoni, J., & Miller, J. H. (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica, 70*(2), 737-753.

Baldwin, J. M. (1896). Physical and Social Heredity. *American Naturalist, 30*, 422-428.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, N.J: Prentice Hall.

Barr, A. (2004). Kinship, Familiarity, and Trust. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence in 15 small-scale societies* (pp. 305-334). London: Oxford University Press.

Barry, H., Child, I. L., & Bacon, M. K. (1959). Relation of Child Training to Subsistence Economy. *American Anthropologist, 61*, 51-63.

Blurton Jones, N. (1984). Selfish origin of human food sharing: Tolerated theft. *Ethology & Sociobiology, 5*, 1-3.

Blurton Jones, N. (1987). Tolerated theft: Suggestions about the ecology and evolution of sharing, hoarding, and scrounging. *Social Science Information, 26*, 31-54.

Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions. *Journal Economic Literature, XXXVI*, 75-111.

Bowles, S., & Gintis, H. (2004). The Origins of Human Cooperation. In P. Hammerstein (Ed.), *Genetic and Cultural Origins of Cooperation*. Cambridge, MA: MIT Press.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *PNAS, 100*(6), 3531-3535.

Boyd, R., & Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethnology and Sociobiology, 13*, 171-195.

Boyd, R., & Richerson, P. (1996). Why Culture is Common, but Cultural Evolution is Rare. *Proceedings of the British Academy, 88*, 77-93.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Bryan, J. H. (1971). Model Affect and Children's Imitative Altruism. *Child Development, 42*(6), 2061-2065.

Bryan, J. H., & Walbek, N. H. (1970). Preaching and Practicing Generosity: Children's Actions and Reactions. *Child Development, 41*(2), 329-353.

Camerer, C. (1996). Rules for Experimenting in Psychology and Economics, and Why They Differ. In W. Guth & E. van Damme (Eds.), *Essays in Honor of Reinhard Selten*. Berlin: Springer-Verlag.

Camerer, C. (2003). *Behavior Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

Camerer, C., & Hogarth, R. (1999). The effects of financial incentives in experiments: A review of capital-labor-production framework. *Cal Tech Social Science Working Paper, 1059*.

Camerer, C., & Weber, M. W. (1992). Recent developments in modelling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty, 5*, 325-370.

Cameron, L. (1999). Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry, 37*(1), 47-59.

Campbell, D. T. (1965). Variation and selective retention in socio-cultural evolution. In H. R. Barringer & G. I. Glanksten & R. W. Mack (Eds.), *Social Change in Developing Areas: A Reinterpretation of Evolutionary Theory* (1965 ed., pp. 19-49). Cambridge, Massachusetts: Schenkman Publishing Company.

Caporael, L., Dawes, R., Orbell, J., & van de Kragt, A. (1989). Selfishness Examined: Cooperation in the Absence of Egoistic Incentives. *Behavioral & Brain Sciences, 12*, 683-697.

Carpenter, J., Burks, S., & Verhoogen, E. (2002). *Comparing Students to Workers: The Effects of Stakes, Social Framing, and Demographics on Bargaining Outcomes*.Unpublished manuscript.

Cavalli-Sforza, L. L., & Feldman, M. (1981). *Cultural Transmission and Evolution*. Princeton: Princeton University Press.

Charness, G., & Rabin, M. (2000). Social Preferences: Some Simple Tests of a New Model.

Cummings, R., & Ferraro, P. (2002). *Inter-cultural Discrimination In The Ulimatum Game: Ethnic Bias and Statistical Discrimination*.Unpublished manuscript, Atlanta.

de Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The Neural Basis of Altruistic Punishment. *305*, 1254-1258.

Durham, W. H. (1991). *Coevolution : genes, culture, and human diversity*. Stanford, Calif: Stanford University Press.

Elliot, R., & Vasta, R. (1970). The modeling of sharing: Effects associated with vicarious reinforcement, symbolization, age, and generalization. *Journal of Experimental Child Psychology, 10*, 8-15.

Ensminger, J. (2004). Market Integration and Fairness: Evidence from Ultimatum, Dictator and Public Goods Experiments in East Africa. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 356-381). New York: Oxford University Press.

Fehr, E., Fischbacher, U., & Gaechter, S. (2002). Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. *Human Nature, 13*, 1-25.

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review, 90*(4), 980-995.

Fehr, E., & Gächter, S. (2002). Altruistic Punishment in Humans. *Nature, 415*, 137-140.

Fehr, E., & Schmidt, K. (1998). *A Theory of Fairness, Competition, and Cooperation*.

Fischbacher, U., Fong, C. M., & Fehr, E. (2002). *Fairness and the Power of Competition*.Unpublished manuscript, University of Zurich.

Fudenberg, D., & Levine, D. (1998). *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Gil-White, F. (2004). Ultimatum game with an ethnicity manipulation: results from Bulgan Cum, Mongolia. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 260-304). New York: Oxford University Press.

Gintis, H. (2000). *Game Theory Evolving*. Princeton: Princeton University Press.

Gintis, H. (2003a). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of Theoretical Biology, 220*(4), 407-418.

Gintis, H. (2003b). Solving the puzzle of prosociality. *Rationality and Society, 15*(2), 155-187.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (Eds.). (2004). *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life*. Cambridge: MIT Press.

Grusec, J. E. (1971). Power and the Internalization of Self-Denial. *Child Development, 42*(1), 93-105.

Grusec, J. E., & Kuczynski, L. (1997). *Parneting and Children's Internalization of Values: A Handbook of Contemporary Theory*. New York: John Wily & Sons.

Grusec, J. E., Saas-Kortsaak, P., & Simutis, Z. M. (1978). The Role of Example and Moral Exhortation in the Training of Altruism. *Child Development, 49*(3), 920-923.

Gurven, M. (2004). Does market Exposure Affect Economic Game Behavior? The Ultimatum Game and the Public Goods Game among the Tsimane of Bolivia. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 194-231). Oxford: Oxford University Press.

Harbaugh, W. T., & Krause, K. (2000). Children's Altruism in Public Goods and Dictator Experiments. *Economic Inquiry, 38*(1), 95-109.

Harbaugh, W. T., Krause, K., & Liday, S. G. (2002). *Bargaining by Children*.Unpublished manuscript.

Hayashi, N., Ostrom, E., Walker, J., & Yamagishi, T. (1999). Reciprocity, Trust, and the Sense of Control. *Rationality and Society, 11*(1), 27-46.

Henrich, J. (2000). Does Culture Matter in Economic Behavior: Ultimatum Game Bargaining Among the Machiguenga. *American Economic Review, 90*(4), 973-980.

Henrich, J. (2003). *A Cultural Species*.Unpublished manuscript.

Henrich, J., & Boyd, R. (2001). Why People Punish Defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology, 208*, 79-89.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (Eds.). (2004). *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies*. Oxford: Oxford University Press.

Henrich, J., & Gil-White, F. (2001). The Evolution of Prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3), 165-196.

Henrich, J., & McElreath, R. (2002). Are Peasants Risk-Averse Decision Makers? *Current Anthropology, 43*(1), 172-181.

Henrich, J., & Smith, N. (2004). Comparative experimental evidence from Peru, Chile and the U.S. shows substantial variation among social groups. In J. Henrich & R. Boyd & S. Bowles & H. Gintis & E. Fehr & C. Camerer (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 125-167). Oxford: Oxford University Press.

Henrich, N. S., & Henrich, J. (2004). *The Cultural and Evolutionary Origins of Cooperation: Kinship, Reciprocity, Social Norms, and Ethnicity among the Chaldeans of metro Detroit*.Unpublished manuscript, Atlanta.

Hill, E., & Sally, D. (2004). *Dilemmas and Bargains: Autism, Theory of Mind, Cooperation and Fairness*.Unpublished manuscript.

Hill, K., & Gurven, M. (2004). Economic experiments to examine fairness and cooperation among the Ache Indians of Paraguay. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & H. Gintis & E. Fehr (Eds.), *Foundations in Human Sociality: Experiments and Ethnography in 15 small-scale societies* (pp. 382-412). London: Oxford University Press.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games. *Game and Economic Behavior, 7*, 346-380.

Hoffman, E., McCabe, K., & Smith, V. (1998). Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology. *Economic Inquiry, 36*, 335-352.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1996). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory, 25*(3), 289-301.

Johnson, A. (2003). *Families of the Forest: Matsigenka Indians of the Peruvian Amazon*. Berkeley: University of California.

Johnson, A., & Earle, T. (2000). *The Evolution of Human Societies* (Second ed.). Stanford: Stanford University Press.

Kachelmeier, S., & Shehata, M. (1997). Internal auditing and voluntary cooperation in firms: a cross-cultural experiment. *The Accounting Review, 72*(3), 407-431.

Kagel, J. H., & Roth, A. E. (1995). *The handbook of experimental economics*. Princeton, N.J: Princeton University Press.

Kaplan, H., & Hill, K. (1985). Food Sharing Among Ache Foragers: tests of explanatory hypotheses. *Current Anthropology, 26*, 223-245.

Knez, M. J., & Camerer, C. F. (1995). Outside options and social comparison in three-player ultimatum game experiments. *Games and Economic Behavior, 10*(1), 65-94.

Kohn, M. (1990). Position in the Class Structure and Psychological Functioning in the U.S., Japan, and Poland. *American Journal of Sociology, 95*, 964-1008.

Larrick, R. P. , & Blount, S. (1997). The claiming effect: why players are more generous in social dilemmas than in ultimatum games. *Journal of Personality and Social Psychology, 72*(4), 810-825.

Ledyard, J. O. (1995). Public goods: A survey of experimental research, *The Handbook of Experimental Economics* (pp. 111-194). Princeton, New Jersey: Princeton University Press.

List, J. A., & Cherry, T. L. (2000). Learning to accept in ultimatum games: Evidence form an experimental design that generates low offers. *Experimental Economics, 3*, 11-31.

Loewenstein, G. (1999). Experimental Economics form the Vantage Point of View of Behavioral Economics. *Economic Journal, 109*, F25-F34.

Marlowe, F. (2004). Dictators and ultimatums in an egalitarian society of hunter-gatherers, the Hadza of Tanzania. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 168-193). Oxford: Oxford University Press.

Marlowe, F. W. (2004). What Explains Hadza Food Sharing. *Research in Economic Anthropology: Aspects of Human Behavioral Ecology, 23*, 67-86.

McElreath, R., Boyd, R., & Richerson, P. (2003). Shared norms can lead to the evolution of ethnic markers. *Current Anthropology, 44*(1), 123-129.

Midlarsky, E., & Bryan, J. H. (1972). Affect expressions and children's imitative altruism. *Journal of Experimental Child Psychology, 6*, 195-203.

Mischel, W., & Liebert, R. M. (1966). Effects of discrepancies between observed and imposed reward criteria on their acquisition and transmission. *Journal of Personality and Social Psychology, 3*, 45-53.

Murnighan, J. K., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology, 19*(415-45).

Parsons, T. (1967). *Sociological Theory and Modern Society*. New York: Free Press.

Patton, J. Q. (2004). Coalitional Effects on Reciprocal Fairness in the Ultimatum Game: a Case from the Ecuadorian Amazon. In J. Henrich & R. Boyd & S. Bowles & C.

Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 96-124). Oxford: Oxford University Press.

Pillutla, M. M., & Chen, X.-P. (1999). Social Norms and Cooperation in Social Dilemmas: The Effects of Context and Feedback. *Organizational Behavior and Human Decision Processes, 78*(2), 81-103.

Presbie, R. J., & Coiteux, P. F. (1971). Learning to Be Generous or Stingy: Imitation of Sharing Behavior as a Function of Model Generosity and Vicarious Reinforcement. *Child Development, 42*(4), 1033-1038.

Pulliam, H. R., & Dunford, C. (1980). *Programmed to Learn: An Essay on the Evolution of Culture*. New York: Columbia University Press.

Rice, M. E., & Grusec, J. E. (1975). Saying and doing: Effects of observer performance. *Journal of Personality and Social Psychology, 32*, 584-593.

Richerson, P., & Boyd, R. (2000). Complex Societies: the evolutionary dynamics of a crude superorganism. *Human Nature, 10*, 253-289.

Rosenthal, T. L., & Zimmerman, B. J. (1978). Social learning and cognition. New York: Academic Press.

Ross, L., & Ward, A. (1996). Naive Realism: Implications for Social Conflict and Misunderstanding. In T. Brown & E. Reed & E. Turiel (Eds.), *Values and Knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Roth, A. E. (1995). Bargaining Experiments. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 253-248). Princeton: Princeton University Press.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittburgh and Tokyo: An Experimental Study. *American Economic Review, 81*(5), 1068-1095.

Rushton, J. P. (1975). Generosity in children: Immediate and long term effects of modeling, preaching, and moral judgement. *Journal of Personality and Social Psychology, 31*, 459-466.

Sally, D. (1995). Conversation and Cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*, 58-92.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science, 300*, 1755-1758.

Sherif, M., I:. (1937). An Experimental Approach to the Study of Attitudes. *Sociometry*(I), 90-98.

Slonim, R., & Roth, A. E. (1998). Learning in high stakes ultimatum games: an experiment in the slovak republic. *Econometrica, 66*(3), 569-596.

Smith, N. (2003). *Are Social Preferences Context Dependent?*Unpublished manuscript.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (in press). Understanding and Sharing Intentions: The Origins of Cultural Cognition. *Behavior and Brain Sciences*.

Tracer, D. (2003). Selfishness and Fairness in Economic and Evolutionary Perspective: An Experimental Economic Study in Papua New Guinea. *Current Anthropology, 44*(3), 432-438.

Tracer, D. (2004). Market Integration, Reciprocity, and Fairness in Rural Papua New Guinea: Results from Two-Village Ultimatum Game Experiments. In J. Henrich & R. Boyd & S. Bowles & C. Camerer & E. Fehr & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from 15 Small-Scale Societies* (pp. 232-259). Oxford: Oxford University Press.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty, 5*, 297-323.

Weibull, J. W. (1995). *Evolutionary game theory*. Cambridge, Mass: MIT Press.

Woodburn, J. (1968). An introduction to Hadza ecology. In R. B. Lee & I. DeVore (Eds.), *Man the Hunter* (pp. 49-55). Chicago: Aldine.

Young, H. P. (1998). *Individual strategy and social structure : an evolutionary theory of institutions*. Princeton, N.J: Princeton University Press.

Endnotes

[1] We extend this axiom to cover cases in which individuals maximize the *expected* utility of their material gains to address the question of risk aversion, but use this simpler formulation otherwise.
[2] Most of this group-level variation is not likely to be explained by differences in sample size between our efforts and those of laboratory experimentalists. First, our experiments used mostly sample sizes on a par with, or larger than, university-based experiments. The robust UG pattern that motivated us is based on numerous samples of 25 to 30 (pairs). For example, Roth et. al. (1991)'s four-country study used samples of 27, 29, 30 and 30. Comparably, the Machiguenga, Hadza, Mapuche and Tsimane used 21, 55, 34 and 70. Overall, our mean sample size was 38, compared to 29 for Roth *et. al.* Second, the regressions on UG offer shown below explain a substantial portion of the between-group variation (which is unlikely to arise via sample variation). Third, we compared this standard regression to a weighted regression (using $1/\sqrt{n}$ as the weight) and found little difference in the results—which shows that the sample size variation is likely not

having important effects. Fourth, we regressed sample size on the groups' deviations from the overall mean (across groups) and found no significant relationship ($p = 0.41$).

[3] If more than one mode is listed, the first number is the most popular offer and the second number is the second most popular, etc. The percent in parentheses is the fraction of the sample at the mode(s). For example, for the Machiguenga 72% of the sample offered either 0.15 or 0.25.

[4]This is the frequency of rejections for offers equal to or less than 20% of the pie.

[5]In Lamalera, Alvard used packs of cigarettes instead of money to avoid the appearance of gambling. Cigarettes can be exchanged for goods/favors.

[6]Instead of giving responders the actual offers, Alvard gave 20 'sham' offers that range from 10% to 50% (mean sham offer = 30%). These are response frequencies to sham offers.

[7] Patton randomly paired Quichua and Achuar players, and as a result there were 14 Achuar proposers and 15 Achuar responders.

[8] Marginal per capita return

[9] Stakes sizes are standardized to one-day wage in the local market, so this column is the endowment received by each player divided by one-day's wage.

[10] The percent in parentheses is the total proportion of the sample at the mode.

[11] Both the experimenters and protocols were identical between Michigan and the Machiguenga (Henrich & Smith, 2004). Comparing the distributions yields a p-value of $p = 0.05$ using the Epps-Singleton test.

[12] Both the experimenters and protocols were identical between the Mapuche and Huinca (Henrich & Smith, 2004). An Epps-Singleton test for a difference between the distributions yields $p = 0.09$.

[13] The 2D intervals are calculated using the following procedure: If the sample has n data points, we create a randomized "bootstrap" sample by sampling n times from the offer distribution *with replacement*. For each randomly sampled offer, we randomly sample a rejection (for example, if we sample an offer of 40%, and two out of three 40% offers were rejected, we sample whether an acceptance or rejection occurred with probability 2/3). This gives a single "pseudo-sample" of n offers and an associated rejection profile of 0's or 1's for each offer. We then use the rejection profile to estimate an IMO (explained in Appendix of Henrich et al., 2004). This single resampling produces a mean offer and IMO. This procedure is repeated 1000 times. Each repetition generates a mean offer/IMO pair. The 2D intervals draw an ellipse around the 900 pseudo-samples, out of the 1000 samples, which are closest to the mean (i.e., the smallest circle which includes all 900 pseudo-sampled (mean offer, IMO) pairs. Small samples generate large confidence intervals because there is a large chance that the pseudo-sample of n draws, made with replacement, is very different than the mean of the actual sample.

[14] A simple measure of our confidence that the average offer is above the estimated IMO is the percentage of re-sampled points that lie below the 45-degree unity line (this is an exact numerical measure of "how much" of the ellipse crosses right and below the 45-degree line). These percentages are 13.7% (Pittsburgh), 0.0% (Achuar), 0.0% (Shona), 58.9% (Sangu farmers), 0.0% (Sangu herders), 1.5% (Mapuche), 1.2% (Machiguenga), 25.5% (Hadza), and 0.0% (Orma). (These figures do not match up perfectly with the visual impression from Figures 4a-b because the ellipses enclose the *tightest* cluster of 900 points, so the portion of the ellipses that overlap the line may actually contain no simulated observations, or may contain a higher density of simulated observations across the 45-degree line). Note that the only group for which this percentage is above half is the Sangu farmers. Even the Pittsburgh (student) offers, which are widely interpreted as consistent with expected income maximization (i.e., average offers are around the IMO; see Roth et al, 1991), are shown to be too high to be consistent with expected income maximization.

The ellipses are flat and elongated because we are much less confident about the true IMO's in each group than we are about the mean offers. This is a reflection of the fact that small statistical changes in the rejections lead to large differences in our estimates of the IMO's. Since rejections may be the tail that wags the dog of proposer offers, our low confidence in what the true IMO's are is a reminder that better methods for measuring what people are likely to reject are needed.

[15] An individual for whom $\rho < 1$ is risk averse, $\rho = 1$ is risk neutral, and $\rho > 1$ is risk preferring. We calculated the values of ρ for which the observed mean offer maximized the expected utility of the proposers, where the expectation is taken over all possible offers and the estimated likelihood of their being rejected. See Appendix of Henrich, et al. (2004) for details on this calculation.

[16] Because the numbers of rejections are small, some of our estimates of risk aversion are imprecise. Accordingly, one concern is that more reasonable estimates of risk aversion might fit the data nearly as well as the best fit. To test for this possibility, we computed the difference between the best-fit value of r

and 0.81, the value estimated by Tversky and Kahneman (1992) from laboratory data on risky decision making. The differences were small for some data sets and quite large for others. In addition, there is a positive but non-significant correlation between the deviation of observed behavior from the IMO and this measure of the precision of the *r* estimate. Thus, it seems unlikely that risk aversion is an important explanation of our observations.

[17] Among non-student adults in industrialized societies, DG offers are higher, with means between 40 and 50 percent, and modes at 50% (Carpenter et al., 2002; Henrich & Henrich, 2004: Chapter 6).

[18] Among the Shona, Barr (2004) used instructions written in such a way as to keep it ambiguous with regard to whom the money is allocated.

[19] Since completing this project, our research team has decided to avoid any use of deception in future work. We also hope to set this as the standard for experimental work in anthropology.

[20] Of course, some variations might matter a lot in some places, but not in others. This kind of culture-method interaction is in itself an important kind of cultural variation.

[21] It is important to distinguish between classes of games in assessing the impact of methodological variables. Many of the largest effects of methodological and contextual variables have been observed in DGs rather than UGs (e.g., Camerer, 2003: Chapter 2; Hoffman et al., 1998). This is not surprising since the Dictator Game is a "weak situation." Absent a strong social norm or strategic forces constraining how much to give, methodological and contextual variables have a fighting chance to have a large impact. UG offers, in contrast, are strategically constrained by the possibility of rejection—that is, a wide range of rejection frequency curves will lead to a narrow range of optimal offers. As a result, we should expect less empirical variation in UG's than in DG's. Thus, one cannot simply say "context matters a lot" without referring to specific games.

[22] Relative wealth was measured by the in-group percentile ranking of each individual, with the measure of individual wealth varying among groups: for the Orma and Mapuche we used the total cash value of livestock, while among the Au, Gnau and Machiguenga we used total cash cropping land. In the UG, estimates using relative wealth were only for seven groups.

[23] The original MacArthur-funded proposal is available at http://www.hss.caltech.edu/roots-of-sociality/phase-i/.

[24] Abigail Barr suggested this procedure.

[25] Three exercises were performed to test robustness. First, since the sample sizes vary across groups by a factor of almost 10, it is possible the results are disproportionately influenced by groups with small samples. To correct for this, we ran weighted least squares in which observations were weighted by $1/\sqrt{n}$. This gives univariate standardized coefficients of 0.61 ($t = 3.80$, $p < 0.01$) for PC and 0.41 ($t = 2.28$, $p < 0.05$) for MI, close to those from ordinary least squares in Table 5. Second, we reran the (univariate) regressions switching every pair of adjacent expressed ranks in the variables PC and MI, one pair at a time. For example, the societies ranked 1 and 2 were artificially re-ranked 2 and 1, respectively, then the regression was re-estimated using the switched ranks. This comparison tells us how misleading our conclusions would be if the ranks were really 2 and 1 but were mistakenly switched. For PC, this procedure gave standardized univariate values of β_{PC} ranging from 0.53-0.66, with t-statistics from 3.0-4.5 (all $p < 0.01$). For MI, the corresponding estimates range from 0.37- 0.45 with *t*-statistics from 2.0-2.6 (all $p < .05$ one-tailed). These results mean that even if small mistakes were made in ranking groups on PC and MI, the same results are derived as if the mistakes were not made. The third robustness check added quadratic and cubic terms (e.g., MI^2 and MI^3). This is an omnibus check for a misspecification in which the ordered ranks are mistakenly entered linearly, but identical numerical differences in ranks actually have larger and smaller effects (e.g., the difference between the impacts of rank 1 and 2 may be smaller than between 9 and 10, which can be captured by a quadratic function of the rank). The quadratic and cubic terms actually lower the adjusted R^2 dramatically for MI, and increase it only slightly (from 0.60 to 0.63) for PC, which indicates that squared and cubic terms add no predictive power.

[26] This is true even for situations of *n*-person cooperation, if punishing strategies also exist (Boyd & Richerson, 1992; Henrich & Boyd, 2001).

[27] Hoffman et al (1994) report similar effects of "social distance" and construal in the UG—for example, players offer less (and appear to accept less) when bargaining is described as a seller naming a take-it-or-leave-it price to a buyer, rather than a simple sharing of money.

[28] It is a common misconception that decision-making models rooted in the preferences, beliefs and constraints approach are inconsistent with notions of evolved modularity and domain-specificity. Such models, however, are mute on this debate, and merely provide an tractable approach for describing how situational (e.g., payoff) information is integrated with coevolved motivations. This implies nothing about the cognitive architecture that infers, formulates and/or biases beliefs and preferences, or what kinds of situations activate which human motivations. It is our view that the science of human behavior needs both proximate models that integrate and weight motivations and beliefs, and rich cognitive theories about how information is prioritized and processed.