

Gil-White, F. J. 2003. "Ultimatum game with an ethnicity manipulation: Results from Khovdiin Bulgan Sum, Mongolia," in *Foundations of Human Sociality: Ethnography and Experiments in 15 small-scale societies*. Edited by J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr, and C. Camerer. Oxford: Oxford University Press.

Ultimatum game with an ethnicity manipulation

Results from Khovdiin Bulgan Sum, Mongolia

Francisco J. Gil-White

University of Pennsylvania
3815 Walnut Street
Philadelphia PA 19104-6196

fgil@psych.upenn.edu

<http://www.psych.upenn.edu/~fgil/>

I. Introduction

This chapter reports two Ultimatum Game experiments conducted with Mongols and Kazakhs in Western Mongolia. The first was an exploratory experiment, and the second is the full-fledged study, with the ethnicity manipulation. This manipulation consists of having proposers and responders be of either the same or different ethnicity.

As with the other studies reported in this volume, one important question is: how will people from a small, non-industrialized society—in this case a traditional pastoral-nomadic culture—largely disconnected from modern capitalism and state structures, perform in the Ultimatum game? If culture has an important effect on economic reasoning and decision-making, then perhaps the rather narrow range of variation obtained from Ultimatum games conducted in the West (and heavily westernized settings) has been a product of Western culture, rather than some more general underlying human nature.

The results on this question are paradoxical. Responders were very reluctant to punish low offers, but proposers were nevertheless very careful not to offend them, and made offers well-above the empirically ascertained income-maximizing offer. I will argue that this result, though paradoxical, is consistent with a parallel paradox that Torgguds experience in their daily lives: people are very afraid of reputation loss, but there are no tangible consequences for those perceived to be 'bad people'.

Another important question guiding this research was to investigate whether the behavior of proposers and responders was affected in any way by having the opposing player be a member of a different ethnic group. The UG experiment with an ethnicity manipulation found an *outgroup*, not an *ingroup*,

favoritism bias. This result ought to be taken seriously in the context of the history of explanations that have been offered for intergroup discrimination. For clarity of organization, I discuss the major relevant theories—Social Identity Theory, and Realistic Group Conflict Theory—in the discussion. But I presage here the most important point: it does not appear likely that these theories can easily account for the results found, and so perhaps we should revisit the assumptions that have justified explaining ingroup favoritism in ecologically valid groups (such as ethnies) in terms of the above theories, which relied on experimental setups that may be much too exotic.

II. Life and economy in Bulgan Sum

This section places the present research in economic context. If the cultural/economic environment affects people's ideas of fairness, and the calculation of costs and benefits, then to understand UG performance differences in different localities we need a competent grasp of the economic framework and standard-of-living conditions affecting our respondents.

A. Ethno-geography

Bulgan Sum is a district in the province of Khovd, Republic of Mongolia. The fieldsite lies against the international border with China, across which is the Chinese province of Xinjiang. The district 'center', Bulgan, is a town that serves as a focal point for the nomadic pastoralists that roam the district, and it is also their legitimate point of contact with state structures. The district is divided into *bag* (or 'brigades'), two of which comprise the town itself, while the rest carve up the nomads roughly along ethnic and clan lines. The *bag* are territorial units, in a sense, but the different lands which constitute the 'territory' of a *bag* need not be contiguous, as some of the nomads, and in particular the Bangyaxan clan, with whom I worked, migrate quite far to their summer pastures. In the process, they traverse territory belonging to other clans. The chiefs of the nomadic *bag* are invariably nomads themselves, are members of the relevant communities by descent and upbringing (which communities elect them), and migrate together with the rest. So the *bag* is in this sense a mobile *social* unit, whose minimal administration (in addition to the *bag* chief there is a forestry chief and a vet—all nomads) travels with it. Nevertheless, the *bag* chief is a member of the state structure, a government functionary, though at the lowest-level, and he answers directly to the elected bosses in the district center.

Torguud Mongols and Kazakhs comprise the two main ethnies in Bulgan. Both Torguuds and Kazakhs are represented among the town dwellers in the district center, and among the nomads who roam beyond it. There are about 5000 inhabitants in town, employed in government, petty trade, artisanship, and small-time horticulture and farming. Many are unemployed. Beyond town is the desert-steppe, where another 5000 people eke out a living as nomads.

During the fall, winter, and spring months, the nomads are not too far from town, though some are much farther than others. Members of the Torguud Bangyaxan clan, with whom I worked primarily, herd their animals to the west of town, in one of the two large floodplains of the Bulgan river, a glacial river that descends from the Altai into the Gobi. A few herders (15% at most) are true nomads, migrating throughout the winter season in the low hill-country or in the sands of the true desert. However, even these have a ‘winter’ property, a plot of land to which they own title with full property rights, which is sometimes equipped with a small bunker used for storage, and always with a corral to protect the animals from winter conditions and to store fodder collected in late August/September. These titled pieces of land are very small and, apart from these, the wilderness is ownerless. Most herders spend the winter in the floodplain and do not migrate during this time. In the summer months all nomads move to high ground, in the Altai Mountains, changing their location constantly as they judge the need for better pastures (they may make as many as 10 migrations in a four-month period). The high ground to which the Bangyaxan clan moves in the summer is very green, high-altitude forest-steppe, crisscrossed by innumerable glacial rivers and streams. Although there are a few Kazakh families in what is essentially Bangyaxan territory, on the whole Torguuds and Kazakhs are territorially segregated—the Kazakh herding to the *East* of Bulgan. Even Torguud and Kazakh town-dwellers live in separate parts of town, separated by a no-man’s land about a kilometer wide.

The town center is a collection of dilapidated and decrepit buildings, of which the largest (and most well maintained) is the school. Except for two small, four-story apartment buildings, the school, and the hospital, all buildings are one-story. The other buildings in the town center are mostly small shops and the administrative buildings of the town government. Some live in small mud houses, and a few others in wooden houses left behind by the military when they used to quarter themselves in Bulgan. Many others live in *ger* (yurts), just like the nomads.

In the summer, the main feature of the town is dust—dust everywhere. I hear that there used to be plenty of trees, but illegal cutting for firewood has left the town with the appearance and feel of a ghost town. There are no paved roads anywhere, as indeed there aren’t in most of Mongolia. Even the airstrip is nothing more than a convenient stretch of flat, dusty, and gravelly ground to the north of town. The result of all this is that the frequent West-to East windstorms (which mercifully are absent during the winter) make life in town very uncomfortable, as they kick up awesome sandstorms that make outside activities impossible whenever they rage. Much of town-dweller small-talk revolves around the miseries of dust and sand. Vicious Central Asian mosquitoes occupy much conversation as well. Add to that the considerable summer heat of the Gobi desert, the stench and discomfort of the latrines, and it does not make for a very pleasant summer in town.

In the floodplains of the Bulgan river, life is considerably more pleasant, for the abundant grass that grows there holds down the dust and sand. Mosquitoes

are even more plentiful than in town, but then the nomads are not there for most of the summer, moving to the highland pastures that are cool, beautiful, and free of dust and mosquitoes. There is no need for latrines in such underpopulated surroundings, and everything smells wonderful. It is small wonder that the herders disdain town life. The cost is in distance to the services located in town, such as they are.

The unforgiving winter is hardly much better for town-dwellers. The mosquitoes and sandstorms leave, but snow and bitter cold take their place. Even here, they don't have much of a comparative advantage with the nomads because the nomad winter *ger* is almost as cozy as a town bunker (*baishang*). However, the nomads do have to labor outside significant amounts of time, whereas town dwellers can afford to be inside for the most part.

B. Brief history

For much of the 20th century, Mongolia was a communist country under the direct influence and tutelage of the Soviet Union, stepping in as a 'protector' state to guarantee Mongolia's borders from the encroachments and claims of the Chinese, and later from Japanese imperialism during WWII. Early on, the Bolshevik revolution spilled into Mongolia, and later Mongolia's own 'people's revolution' was engineered by Soviet agents and their puppets on a stunned populace. The excesses of this time were many, including wholesale persecution and slaughter of dissidents, especially members of the feudal Lamaist clergy.

In time, a command economy was instituted where everybody, nomad or town-dweller, was a government employee, but unlike the nomadic peoples of the Soviet Union, Mongolian herders were by-and-large not forcibly sedentized. Rather, the herders were organized to herd animals for the collectives, called *negdel*, which owned the animals and paid them wages. The government distributed medals and bonuses to good workers but effort and talent were not really an avenue to riches. Jobs, pensions, and health-care were subsidized and guaranteed, and most services were delivered through the *negdel*. The subsidies for these services were paid mostly by large Soviet outlays rather than broad-based taxation, which generated very little revenue.

After the democratic revolution in 1990 the command economy collapsed with the disappearance of Soviet subsidies. Livestock was redistributed and the economy privatized. Initially, a few herders tried of their own free will to reorganize in voluntary collectives called *xorshoo*, but after only two or three years, even these experiments were abandoned for full privatization. This of course means that only for the last nine years have Mongolians experienced the connection between work and wealth/security which is characteristic of societies with private property. The herders have readapted remarkably fast, rebuilding kin- and reciprocity-based patterns of association that existed prior to the revolution to substitute for (now absent) state services (Szynkiewicz 1993). Nevertheless, many are understandably nostalgic for the subsidized command economy.

C. Cash economy, development, and market penetration:

This section examines the degree to which this site is different from more industrialized settings. It is important to know the extent to which people in Bulgan use cash (since the experiments involve cash), the degree to which they participate in markets, and the degree to which they have access to services typical of a developed country. The prices of different things, and in particular of hired labor, are also examined.

These days, the country folk in Bulgan Sum don't earn wages and the town folk do—that is, those who still have a job. In town unemployment is supposedly rampant but, except for government jobs, those with an occupation employ themselves, so real unemployment statistics are hard to come by. Since during communism nomads earned cash wages, the economy may have become *less* of a cash economy than it was before. The state has also receded from view considerably; most noticeably for the nomads, in the provisioning of basic services such as health-care.

Among nomads, cash is used but many (perhaps most) transactions are barter (as indeed are many transactions even in town). Cash is used primarily on several yearly trips that nomad women make across the nearby Chinese border to get clothes, flour, and other things. Among themselves, the nomads trade almost nothing, for trade efficiencies are few when everybody has virtually the same way of life. There are two prominent exceptions: Kazakhs tend not to breed as many horses as Mongols, from whom they buy them. The same appears to be true for camels. On the other hand, Kazakhs make horseshoes and knives, which Torguuds buy.

Many townsfolk earn wages or free-lance honoraria, and most petty commerce transactions are for cash. On the whole, however, there is a great deal of friend-to-friend quid-pro-quo and kin-based altruism/reciprocity involving no cash. Some of the most important exchanges of this sort occur between the nomads and townsfolk.

1. Food. The food economy is largely cashless. Important items in the diet such as aaruul (cheese curds) and butter are obtained by townsfolk from herders as gifts or as part of loose and informal reciprocal exchanges. Most townsfolk own livestock which is herded by herder relatives or friends in exchange for housing their children during the school year. In addition townsfolk keep cows or goats for milk, or else they make do with low-prestige *khar tsai* (black tea) rather than *süütei tsai* ('tea with milk' the traditional beverage of Mongols, which contains about 1/3 milk). Some townsfolk plant a few horticultural products such as cabbage and potatoes. Some nomads and townsfolk plant a little rye, but wheat-flour they buy for cash either in Bulgan Sum, or else the cheaper (and lower quality) flour from China. A 25 kg. sack of wheat-flour cost T6000 (≈\$7.50) in 1998.

Other items bought for cash include salt (wildly abundant in the Gobi and brought to town by industrious people with a vehicle) and tea. The latter is the most omnipresent item in the Mongolian diet. Everybody pays cash for this, and it comes either from Russia (high-quality) or from China (low-quality). It is considered expensive item and is sold in bricks of pressed stems and throwaway leaves that typically are not sold in other parts of the world. One brick of tea sells for T1500 (about \$2) and it lasts a family of 5 for about a month and a half. There are also incidental items such as vodka, fruit, candy, cookies, hydrolized protein, chili sauce, sugar, etc.; but none of these are consumed in great quantities, except for vodka, sugar, and candies.

2. Services. A *tandig jolooch* ('a driver you know'), will give you a free ride or merely take gas money if you don't go far. When my friend Baajaa and I drove to the floodplain once in his truck, we transported three people's belongings, but he didn't charge *any* of them for his services, and took gas money from only two because the other one was very poor. Baajaa was going in that direction anyway, and this appears to be the norm: drivers do favors to people who are going their way. There is no regular public transportation, and for long trips one must hire a driver outright. In town people simply walk unspeakable distances, or stay put.

Every week, a small plane from Ulaanbaatar stops first in the provincial capital of Khovd and then lands in the airstrip to the north of town. The reliably awful windstorms may cause delays of a few days, or outright cancellations. The round trip from Ulaanbaatar was, in 1997, a subsidized \$50 for Mongolians and \$350 for the foreigner anthropologist. The alternative is to brave 1500 km of dirt roads for 3 to 4 days (including nighttime driving) in good weather.

About 40 families have a phone—out of about 500 that live in the perimeter to which phone lines reach (for some of these, the phones no longer work). There is only one public phone in town, at the phone company headquarters, and one must pay cash to use it. I made use of it several times, and at worst never saw more than ten or twelve people waiting (often just one or two). Most people have no use for the phone or find it too expensive. They may send letters instead, although this will be expensive too—T220 (about 22c) to Ulaanbaatar. Just sending four-to-six letters will eat up a day's wages for a hired hand (see below).

These days the diesel to run the electricity plant is no longer subsidized by the USSR and must be bought from the Chinese just across the border. It is too expensive. To fill up Baajaa's truck's tank, for example, requires all of US\$25, or about three weeks' salary for a hired hand (see below). The town can afford to run its new Japanese plant only for four hours a day, and only in the winter, when electricity is absolutely indispensable given the short duration of sunlight at this latitude.

Fresh water comes from wells, easily dug since the water table is only a few feet below the surface. Human feces are disposed of in latrines. The nomads, on the other hand, get their water either from wells or from streams, and practice

open-air fecalism. The one service that surprisingly—and thankfully—does exist is garbage collection. The trash is deposited on the sands, downwind, to the east of town.

There is a bank in town which apparently is used by very few people and does nothing more than provide safe-keeping and petty lending, much like a frontier bank in the old American West.

3. Clothing. A prominent item in the cash economy. Nomads buy material for cash and the women confection and tailor their *dels* (long coats). They also make felt ‘socks’ for their winter boots. Aside from this, however, everybody wears Western-style clothing brought from either China (much of it cheap surplus dumping) or Ulaanbaatar and which is sold for cash—e.g., a pair of boots T20,000 (≈\$25), a pair of socks T300, a shirt T1500, and a pair of pants T1500, at 1998 prices and exchange rates.

5. Industry. There is really no primary production of any sort in Bulgan Sum except for the products of nomads and farmers. The only real middleman economy, however, is for products produced elsewhere, such as clothing, fruits, candy, school supplies, etc. It is not a large economy, but it is the largest sector in town by far (perhaps forty tiny stores in town, all of them virtually identical in size and merchandise). There are also people who buy skins and fur from the nomads and sell them elsewhere. In addition there is a very small secondary processing industry that turns grain into flour for local consumption, and another which minimally processes livestock entrails for sending to China where real value is added by turning them into such things as hats which are then sold back to the nomads. I have seen this tripe ‘industry.’ It is a sorry-looking bunker about the size of three typical *baishang* (2-room adobe winter-dwelling) put together, and supposedly employs all of ten people. I suspect the flour ‘industry’ supports the same. That is the extent of industry in Bulgan.

6. Labor for hire. Every once in a while somebody needs an *ajilchin* [hired worker], even among the nomads, and this is perhaps the best measure of how much money, in relative terms, my subjects were playing for in the Ultimatum experiments. Consider that the stake was T8000, equivalent to about \$10 at the time of the experiments.

Ariüngerel (a Torguud nomad) one day hired another herder to do some work for him in exchange for a small goat that he reckons goes for T3000 to T4000. The worker never delivered and Ariüngerel complained loudly to his sympathetic sister-in-law and her children. He expected to get three or four days labor, or else the man should give him back an equivalent goat. On a separate occasion (much later) I learned from him that for a *full-grown* goat (valued at T8000) the *ajilchin* would work for eight days. So his answers on two separate occasions are consistent, and indicate a daily wage rate of T1000. In this reckoning, and comparing it to the US minimum wage, people are playing for the equivalent of \$400. This is obviously a distortion because the price of labor relative to the prices of other things is almost certainly cheaper here than in the US. It nevertheless indicates that people are playing for definitely non-trivial

sums: a week's worth of your labor is a week's worth, in Bulgan Sum and in the US. For a comparison estimate, I asked Kairatpic (a Kazakh nomad), and he told me that a hired hand would get T1500 daily, and this is not too far away from Ariüngerel's estimate.

My friend Baatarcüren has the most beautiful *baishang* that I have seen among Torguud nomads. It is so well-finished one could almost mistake it for a Kazakh *baishang*, which are much nicer given that Kazakhs have a stronger tradition of living in houses, whereas Torguud nomads typically only build bunkers for storage, spending the winters in their *gers*. It is a two-room 7×4m cabin made out of wood (but not logs), built by Aavda, who finished carpentry school. Baatarcüren provided him with all of the materials but helped not at all with construction. It took Aavda 5 days to complete it, and he was working 8:00 am to 8:00 pm days. Baatarcüren gave him T20,000 for the job. This means that Aavda was getting T4000 per day. That seems like an awfully high wage-rate. These high wages may be explained by the fact that carpenters are rare (in Bulgan Sum there are no more than 4). Baatarcüren tells me that they are invariably very well-respected people—Ariüngerel's *ajilchin*, on the other hand, had been hired to do unskilled labor. Besides, Aavda was working 12-hr days, which is unusual. A driver from Bulgan Sum who came to Baatarcüren's looking for his lost cows (a common occurrence), told me that the norm is for a hired hand to work 8-hr days, with a daily wage adjusted for the heaviness of the task but never higher than T3000 and, on average, T1500. In light of all this evidence I settle on a conservative estimate daily wage rate of T1500 which may be inflated given the incentive to do this with the rich foreigner. My hired assistants for the experiments all got paid higher than this: a daily T1600 and, after the devaluation, T2000, but always the equivalent of \$2. They seemed to think that they were being paid handsomely.

III. The experiments

A. Study 1

The first study unintentionally ended up as an exploratory one, for its data were inconclusive due to unforeseen problems with the methodology. Nevertheless, the data are very suggestive, and a review of the evidence obtained, and the lessons learned from it, help increase the confidence for the results in study 2. It was conducted exclusively with Torguud nomads from the Bangyaxan clan, while in their high-altitude summer pastures. Thus, this is a simple UG experiment except for the idiosyncrasies of this field-site and its methodological requirements.

1. Methods

The study population consists of 40 adults ranging in ages from 18 to 56. There were 20 proposers (12 men and 8 women), and 20 responders (9 men and 11 women).

The clan veterinarian, Boldoo, is an intelligent 30-year-old nomad with higher, technical education, and familiar with the notion of an experiment and its methodological requirements. Moreover, he could see why it was interesting. He was hired as an assistant. At the time of this experiment, my Mongolian was not enough for conversational fluency, but it didn't prevent communication or explanation. Boldoo corrected the Mongolian in the draft of my explanations and was instructed to intervene only if, after failing to communicate a particular point, I asked him to do so. This way I could have the highest degree of control over what was said, especially since my Mongolian was adequate for me to almost always understand Boldoo's interventions.

The dispersed manner in which steppe-nomads live made it impossible to gather all participants for the experiment, so the experimental protocol unfolded as follows:

1. Twenty responders were recruited but the game was not explained to them.
2. For every responder, a Polaroid was obtained.
3. Twenty proposers were recruited and the game was explained to them.
4. Each proposer, after having the game explained to him/her, made a proposal, having been advised that the responder was a person among the 20 photographs I presented, but whose precise identity I could not reveal, as everybody would play anonymously.
5. Following the offer, a Polaroid was taken of every proposer.
6. Responders were assigned randomly to proposers.
7. Another trip was made to each responder's tent.
8. The game was explained to each responder, and the offer made by the proposer was demonstrated, pointing out that the proposer was among the 20 people in the photographs presented (this time, of proposers), but this person's identity would remain anonymous. If the responder accepted the offer, the offered money was immediately apportioned to her/him.
9. A final trip was made to each proposer's tent. The responder's response was conveyed, and any money due was apportioned.

Appendix A has the details of the methodological protocol for the explanations used during proposer recruitment (the explanations to the responder were essentially identical, with a few obvious modifications necessary to make it specific to responders). I was always left alone in the *ger* with the proposer/responder when witnessing their choices, when revealing how much money was due to them and making the payments, and during cross-examinations

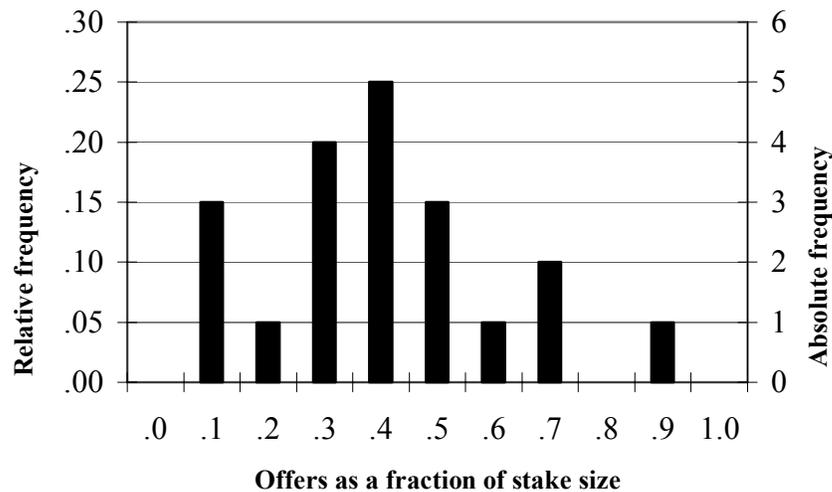
(see below). Each one was assured of confidentiality, and I explained that they could not even be identified by name in my charts for everybody was getting a number.

Proposers were disqualified if they failed repeatedly to understand the experiment, as evidenced by incorrect responses to iterated quizzes after re-explanations. Only one proposer was disqualified in this manner. Unfortunately, initial smugness over the ease with which the experiment would be understood resulted in the following serious methodological flaw: no disqualifying procedure was employed for the responders, as these had the experiment explained to them *after* they had been recruited, and after the proposers had made their offers. This shortcoming cost me two data-points, because two of my responders proved incapable of comprehending the logic of the experiment, and I couldn't use their responses.

After apportioning any money due, I cross-examined participants to examine their reasoning. In particular, I sought answers to the following questions: (1) what was the punishing threshold that they anticipated, and was this answer consistent with their reasoning in the experiment, and with their offers?; (2) what is the lowest amount that the proposers themselves would have accepted, had they been responders?; (3) for those responders who had rejected the offers made to them, why had they?; and finally (4) for those responders who had accepted the offers made to them, was there a level at which they would consider the offer too miserly to be accepted?

2. Results and discussion

Fig. 1. Distribution of offers in Study 1.



There is no difference between males and females (male mean = 0.43, female mean = 0.46; $CF = 4.6, p = .33$). The mode (0.40) and mean (0.44) of this distribution are not too surprising from what we know of experiments carried out in developed settings (see Camerer and Fehr, this volume). However, the variation seems rather wild, especially considering that it includes 4 data points (0.20 of the total) above the 50/50 line, including a very high outlier at 0.90, and also a rather largish spike at 0.10. This was surprising.

There was only one rejection, and it went to a 0.10 offer. This too was surprising, especially given that fully 6 offers were for 0.30 or less, and that even this lonely rejection may not be real. It was given by Boldoo's father, Shavia. Under cross-examination, Shavia told me that he had rejected the offer because he thought it was 'ugly,' and he wanted the proposer to take a loss. This reply pleased my prejudices, but as we departed, an amused Boldoo relayed to me that when his father had stepped out of the *ger* he had told Boldoo that his real reason for refusing the offer was that he could not accept money from a poor student like myself. This charming generosity, from a man whose entire property I could have bought with a little bit of credit card debt, was entirely in character for the late and sorely missed Shavia (he died in 1999)—a wonderful host, father, husband, and friend. I went back and confirmed Boldoo's statement with Shavia, and explained that the money was the university's and it did not affect me whether he took it or not. Had he understood—he informed me when queried—he would have rejected the offer anyway. But I must be skeptical because Boldoo—who had discovered that this is what I wanted to hear—was whispering the answer in his father's ear.

There were other confusions. I undertook a second round of cross-examinations, and the sum total of these investigations yielded even greater uncertainty over the proper interpretation of the experiment, as it became obvious that many participants had not understood the point of it all. Some of this information came to me serendipitously. For example Galcnii Mönxbat's offer of T7500 (only one T500 increment below giving everything away) seemed ridiculously high to me. His brother Anxbat (who had gone before him, and assisted Boldoo and myself in the role-playing part of the explanations), Boldoo, and myself, were all satisfied that he had understood the object of the game. In the cross-examination, I again was not able to uncover evidence that he had misunderstood. The day after, Boldoo and I relaxed in Mönxbat's *ger*, the two of them chatted. Boldoo let out a big laugh and then told me that Mönxbat had not understood the game after all. Mönxbat, he explained, thought he was supposed to get the responder to agree to the partition, but didn't realize that real money was involved and that he could make real money in the experiment. He thought the money aspect of it was hypothetical and had been very surprised when I brought him back his winnings. Thus, he naturally had offered T7500 (0.93) because he reasoned that nobody in his right mind would disagree to such a partition. Had he understood, he said, he would have offered half. His wife, Oyungerel, had misunderstood the experiment in exactly the same way (though her offer was in fact 0.50). At least 3 other people made the same mistake.

Again, this was initially very surprising. My explanation protocol was careful, step-wise, gently paced, and heavily demonstrative. Moreover, it included a test to see whether participants had understood. Thus, it was at first bewildering and discouraging to find in the second round of cross-examinations that several participants still had not understood (even though they had passed my test). My methods, careful as they were, could not anticipate the great cultural gulf separating me from my respondents, for I could not have imagined some of the hypotheses they made concerning the object of the game. Nor could I have anticipated some of their scruples given that some incorrectly believed the money to be mine.

The cross-examination of the responders added to my growing skepticism. Most responders didn't seem to have a punishing threshold, and, when they did, it was extremely low. It was not consistent with proposer behavior that responders would accept any offer, and neither was the strong reluctance I found in responders to characterize low offers as punishable or particularly miserly. These data from the responders quite directly contradicted what the proposers said they would have done had they been responders themselves. Most proposers indicated that they would have punished low offers (though what they thought was a punishable offer was not always the offer immediately below the one they themselves made).

All of this led me to doubt that responders had truly understood the experiment. I hypothesized that, despite my precautions, they might not have understood that the proposer was *required* to make an offer in order to play the game. If so, they might have regarded any offer as a gift, and missed the point of the game, which lies in the power of the responder, and the respect or disrespect for this power implied in the proposer's offer. A corollary hypothesis was that the logic of the game might be easier to grasp for those who are socialized into the proposer role, and this might explain the discrepancy between responder behavior and the answers proposers gave as hypothetical responders. Study 2 was designed with an eye to improving the methodology in light of all of the misunderstandings, to resolve these troubling theoretical questions, and to test for possible ethnicity effects in ultimatum game decisions.

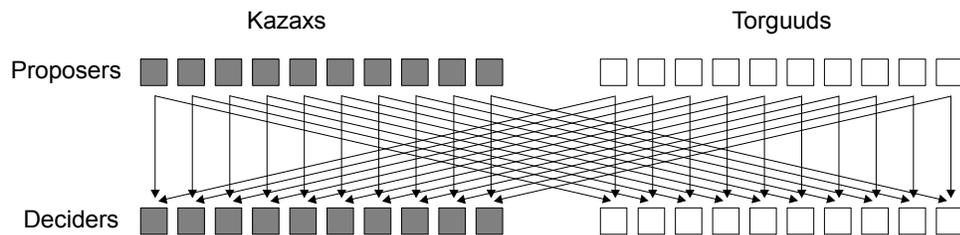
B. Study 2

As originally intended, Study 1 would provide an ingroup baseline (in Study 1 all of the participants are Torguud Mongols) against which to compare the results in Study 2, where Mongols and Kazakhs would be playing against each other. In this manner I would test for an ethnicity effect. However, because of the problems listed above, I was not confident that I could interpret Study 1. Thus, I redesigned Study 2 so that I could simultaneously fix the methodological problems in Study 1, *and* obtain both in-group and inter-group performance data.

1. Methods

The study population consists of 40 adults, ages 20 to 50. They were not nomads but town-dwellers (except for a few nomads among the Kazakh). Half of them are Torguud and half Kazakh. Of the 20 Torguuds, 10 are proposers and 10 responders, and likewise for the Kazakhs. The ratio of males to females per ethnic per role, in both cases, is 60/40. Each proposer made two offers, one to a Kazakh responder, and one to a Torguud responder, anonymously. Each responder thus had to make two decisions, one for the offer a Kazakh made, and one for the offer a Torguud made. The diagram below illustrates the two offers every proposer must make, as well as the two decisions facing every responder. Proposers and responders were assigned at random.

Fig. 2. Structure of offers and responses



The following steps summarize the experimental protocol.

1. Twenty responders (ten from each ethnic) were recruited and the game was explained to them. They were tested, and those who failed were disqualified.
2. For every responder, a Polaroid was obtained.
3. As in step 1 for the proposers.
4. Each proposer, after having the game explained to him/her, and passing the test, made a proposal. Each proposer was told that the responder was among those portrayed in 20 photographs which I presented, but whose precise identity I could not reveal, as everybody would play anonymously.
5. Following the offer, I explained that there was a second responder with whom the proposer would play. I was very careful to point out that this was a completely new game, played with a fresh batch of T8000 *tögrög*. A new batch of 20 photographs was presented (half of the participants made an ingroup offer first, and the other half made an outgroup offer first).
6. Following the second offer, a Polaroid was taken of every proposer.
7. Responders were assigned randomly to proposers.
8. Another trip was made to each responder's dwelling.

9. The object of the game was quickly reviewed. The offer made by the proposer was explained, pointing out that the proposer was among the 20 people in the photographs presented (this time, of proposers), but this person's identity would remain anonymous. If the responder accepted the offer, the offered money was immediately apportioned to her/him.
10. A final trip was made to each proposer's tent. The responders' responses were conveyed, and any money due was apportioned.

For the Polaroids I always insisted on ethnic headwear. This is more predictive for the Kazakhs, as no Torguud wears the Kazakh *kapes*, but some young Kazakhs sometimes wear the commercially bought hats that the Torguuds favor. However, the ethnic manipulations worked every time—every single participant was aware that one set of photos consisted of exclusively Torguud individuals, whereas the other consisted of exclusively Kazakh individuals.

The new recruitment procedure included the one used earlier *in full* (Appendix A), including the test contained therein, followed by a new, second test, which tried to search exhaustively for any of the likely misunderstandings (Appendix B). Every potential participant had to pass the new test before being admitted into the experiment. The earlier protocol was incorporated with only two differences: (1) I now used two assistants, so that I could play myself in the explanatory pantomime and preempt confusion; (2) no participant was paid merely for participating. This latter change was due to the considerable unease and consequent distraction (not to mention extra time devoted to explaining experimental custom in the US) caused by their bewilderment at my trying to pay them for agreeing to humor me while I asked them some questions—with which they might make money to boot! It was obvious that their bewilderment occupied their minds considerably and interfered with their comprehension of the explanations surrounding the object of the game (for example, they were often confused about whether they could keep the participation payment, or whether it was part of the money for play). The effort to treat these participants with no less respect and consideration than is due to Western subjects is laudable, but this must include sensitivity to the cultural situation. If importing a given cultural value of ours to the field distracts and unsettles the participants, in addition to imperiling the experiment, its wisdom in the field setting ought to be reconsidered.

In Study 1 I found little evidence of a punishing threshold, whether in performance or in answers to hypothetical questions. In order to test whether this had been a problem of comprehension or a true populational attribute, I used the following procedure when presenting the offer to the responder:

1. On my chart (which only I can see) I pretend to search for the line that corresponds to this person's anonymous proposer.
2. I pretend to find it and I demonstrate for the responder the lowest possible offer, saying "This is the offer that the proposer made. Do you accept or reject?"

3. After the responder gives the answer, I write it down in my chart, and, as I do, I pretend to notice that I made a mistake, claiming to have read a different line which corresponds to somebody else's proposer.
4. I then demonstrate the real proposal, and record the responder's answer. Naturally, any money due to them came from how they responded to the real offer.

In this manner I obtained very good data for reactions to the lowest possible offers. Half of all responders saw the outgroup offer first, whereas the other half saw the ingroup offer first. The fake offer was always presented as coming from whichever proposer came first. My only defense for thus breaking with the conventional norms of experimental economics is that I was neither knowledgeable enough nor creative enough, at the time, to come up with the titration method, where subjects commit themselves in advance—before seeing the offer—to the quantities that they will either accept or reject. Faced with the decision of keeping to the norms or coming back without having answered the question as to whether responders were likely or not to punish low offers, I opted for breaking the norms.

2. Results

The considerable extra work involved (not least because the test led to many disqualifications, which multiplied the number of times the exhausting explanatory routine had to be repeated) seems to have paid off. Out of 20 proposers, 13 showed an 'excellent' performance in the test, which means they understood all the smallest nuances (e.g. guessed right the first time whose money they were playing with and so forth), even though these were not overly emphasized in my explanatory routine. Another 5 got a 'pass,' which means they made one or two mistakes, but not on the basics, and had no trouble understanding the right answers when I explained them. Only two people were passed with some reservations. The numbers for responders are virtually identical.

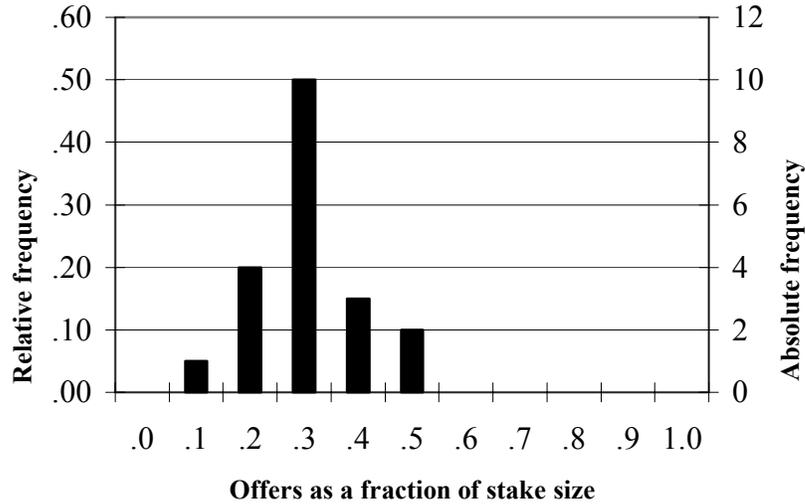
Is the increased confidence in participant comprehension reflected in corresponding changes in the data? It is for the distribution of offers. The table below shows the means of Torguud offers to an ingroup and an outgroup responder, and the corresponding Kazakh means (each sample n=10), with the Epps-Singleton test results comparing the Torguud and Kazakh distributions.

Table 1. Torguud and Kazakh offers in Study 2

	By Torguuds	By Kazakhs	CF	<i>p</i>
ingroup offer mean	0.350	0.356	2.63	0.62
outgroup offer mean	0.418	0.431	6.08	0.19

Since the Torguud and Kazakh offers do not appear statistically different, I group the combined ingroup offers and compare them to the offers in Study 1 to see if the new methodology has affected the results.

Fig. 3. Combined Kazakh and Torguud *ingroup* offers in Study 2.



Compared to offers in Study 1 (see **fig. 1**), the offers are lower. This makes sense given the kinds of misunderstandings that resulted in Study 1: that the money was not for real, that the object of the game was just to get the responder to agree, that the money was mine, etc. When the game is properly understood as one where real money can be made, and where making money will not harm the researcher economically, people who wish to make money will tend to make smaller offers. In this second study there was only one offer above 0.5, and it was only marginally so (0.56), whereas in the first study there were 4 such offers, including one at 0.9. This is a big part of the difference.

The Epps-Singleton test adds weight to the conclusion that the distributions are different, as shown in **Table 2**.

Table 2. Comparing Study 1 and Study 2* offers

	Study 1	Study 2	CF	P
mean	0.44	0.35	8.27	0.08

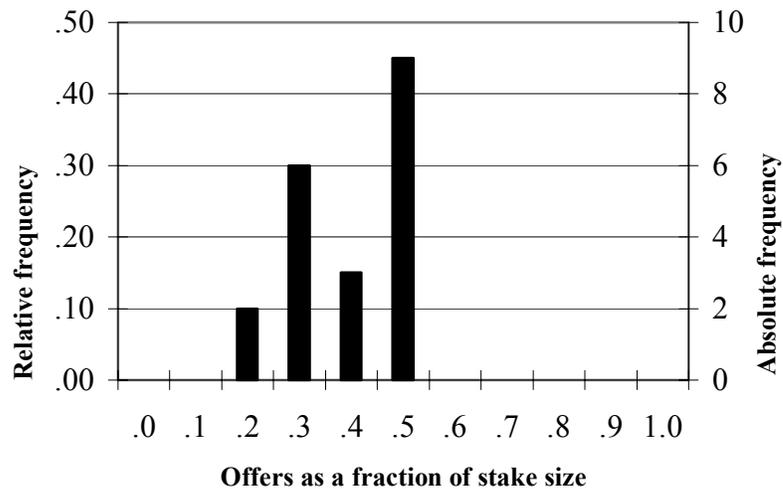
*The Study 2 sample is of the combined Kazakh and Torguud *ingroup* offers

The result falls short of the conventional threshold for statistical significance concerning the apparent difference between the two distributions, but we have data in excess of the statistical test. We have systematic interview and fortuitous ethnographic data concerning people’s misunderstandings in Study 1

(and these data *underestimate* the potential misunderstandings). We also have a thorough test and rigorous disqualification procedure in Study 2 that rules out such misunderstandings there. Moreover, the idea that the distributions would be different in the direction found makes intuitive sense if the methodology was indeed successful at removing such misunderstandings.

Next, does facing an ingroup or an outgroup responder affect the size of the offers? First, I group together the outgroup offers (Torguud *and* Kazakh) as I did above for the ingroup offers.

Fig. 4. Combined Kazakh and Torguud *outgroup* offers in Study 2.



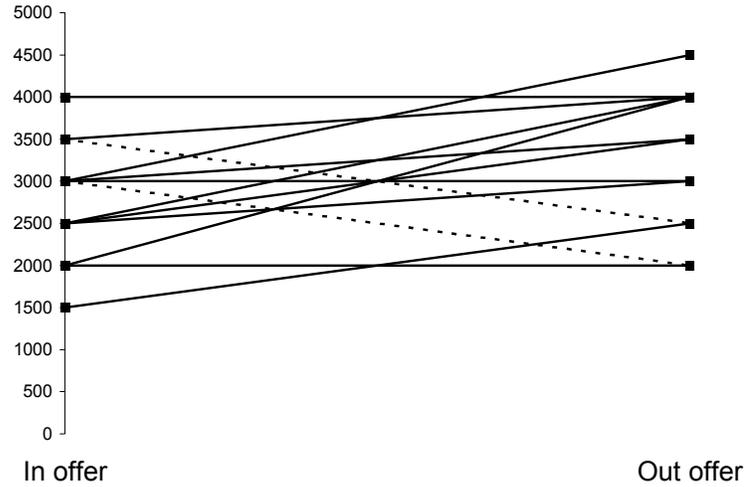
Comparing the distributions in **fig. 3** and **fig. 4** shows them to be different.

Table 3. Offers to ingroup vs. to outgroup responder; Paired T-Test

	Ingroup	Outgroup	T	p
offer mean	0.35	0.425	-2.88	0.01

The effect is perhaps better appreciated in the following figure.

Fig. 4. Individual in/outgroup offers.



Only two people gave an ingroup offer higher than their outgroup offer. Everybody else either stayed the same, or gave a higher offer to the outgroup responder (some lines overlap exactly, so the total number of apparent lines is less than 20).

Every responder faced one fake offer (and then two real ones). Half of all responders thought the fake offer came from an ingroup, and the other half from an outgroup, member. The table below summarizes their responses.

Table 4. Responder responses to fake, T500 (lowest possible) offers

	Accept	Reject
Ingroup offers (n=10)	7	3
Outgroup offers (n=10)	7	3

Respondents seem oblivious to whether the low offer is from an ingroup or an outgroup proposer, and a strong majority doesn't think that an offer which is T500, or 0.063 of the total T8000 stake, deserves punishment. The real offers, naturally, were accepted at a higher rate (3 rejections out of 20 real offers). The behavior of the three rejecting responders is summarized thus:

**Table 5. The three responders who rejected real offers
(A = accept; R = reject)**

Responder	Ingroup offer	Outgroup offer
1	0.375 (R)	0.25 (A)
2	0.313 (A)	0.375 (R)
3	0.50 (A)	0.375 (R)

Responder 1 rejected an ingroup offer but accepted a *lower* outgroup offer. Responder 2 by contrast rejected an outgroup offer, but accepted a lower ingroup offer. Responder 3 rejected the low offer and took the high one. So one responder's behavior (3) yields no information about group bias, and the other two cancel each other out. In summary, the data from rejections yielded no evidence of group favoritism, whether ingroup or outgroup. But the failure again to find evidence of ingroup favoritism must be noted.

One thing does emerge very clearly from the rejection data. In Study 1 it seemed to me counterintuitive and skewed that responders should accept anything at all in light of the fact that the offers were, relative to this reluctance to punish, high. I speculated that this might be due to methodological errors in Study 1 that caused responder lack of comprehension (even though I was not able to uncover this clearly in my interviews). Now I believe that this strange result is real: responders will take practically anything. Despite this, the offers, even in Study 2, are relatively high.

C. General discussion

1. Why doesn't responder behavior predict proposer behavior?

It is puzzling that the offers of proposers are so high when the punishing threshold is so low. Mine is not an isolated result. Except for perhaps the Machiguenga (Henrich 2000) the results we have for simple societies suggest that people are, for the most part, not making income-maximizing offers (see Henrich *et al*, this volume). This could be because people have a stronger preference for avoiding losses than they do for maximizing income (in the sense of Prospect Theory). Or it could be because, although people wish to maximize income, they have a poor estimate of where the punishing threshold is. If there is a bias to assume that this threshold is higher than it actually is, offers will seem large relative to empirically determined rejection rates.

At least for the present case, this second hypothesis receives some support from the cross-examination data in the first experiment.¹ I queried proposers, immediately after they had made their offers, on what they thought was the

¹ The protocol in Study 2 was so tiresome that I did away with the very thorough cross-examinations I had used in Study 1, so I don't have this kind of data for Study 2.

punishing threshold for the anonymous responder. This I did with role-playing and demonstration, as I had not yet truly mastered the use of counterfactuals. For every proposer, I would demonstrate the lowest possible offer (T500) on the ground with actual bills and would say, “What do you think the responder thinks? If you make this offer, do you think the responder will accept?” If they said no, I would up the demonstration offer by T500, and ask again—and so on until they said this was an offer the responder would be inclined to accept. Most proposers said that rejections would happen all the way up to and including the offer right below the one they had actually made. In this sense the proposers were consistent even though the individual estimates of punishing thresholds (and corresponding offers) varied. When I used a similar routine to find out what these proposers would have done as responders, the typical answer was that they would have punished offers below the one they had made.

There are two puzzles here. The first is that the proposers overestimate the rejection rate. The second is that they predict they themselves would punish at the rate that they predict others will even though this is almost certainly not true (given that people were assigned to proposer and responder roles at random). I examine each puzzle in turn.

The tendency in proposers to overestimate the rejection rate—i.e. to overestimate the negative consequences of their ‘bad’ behavior, can be reconciled with the ethnographic data. These data revealed that people were quite neurotic about their reputations and feared greatly the possibility of being perceived as a ‘bad person.’ Moreover, they appeared hyperaware (at least by my standards) of the presumed or real moral qualities of individuals, and very prone to speaking of them in moralistic and essentialist terms—for example *sain khün* or ‘good person’; *muu khün* or ‘bad person’; *tsagaan khün* or ‘a person without any malice or guile’ (lit. ‘white person’). I remarked about this to my friend Tsoloo, who grew up in Bulgan Sum as a nomad but has lived in the capital city for a few years. Full of surprise that I understood this, he exclaimed that people in Bulgan Sum were indeed quite concerned that they not be thought of as *muu khün*.

Despite all this, nothing at all seems to befall those with a bad reputation. The data for this comes from people’s introspections, when I ask them what the likely social consequences will be to someone with the reputation for being a ‘bad person.’ To this question people consistently drew a blank, no matter how hard I tried to tease it out of them. I began collecting such interview data early on, as part of a general interest in social punishment, and did not realize its relevance to the experiments until they were concluded.

What I found is that people claim they will not change their behavior towards ‘bad people’ in any way, even when I suggest mild forms of social punishment such as visiting ‘bad people’ less. And I found reluctance to punish even when I asked about admittedly serious transgressions, such as free-riding on the work of others in one’s *khot ail*. People certainly do gossip, but more significant punishments (of the kind that require the punisher to incur some tangible costs to him or herself) are apparently not typically mobilized against social offenders.

My ethnographic observations are consistent with this, and illustrate with a dramatic anecdote. X, an elderly man who lived close by to my first host, H1, was early on described by the latter as a very bad person. This was supposedly a family shortcoming. I was assured by H1 that this was his general reputation in the community. I later heard other people refer to X in negative terms. In the end it became clear that such an opinion of him is not universal but it is also not limited to my initial host's family—although, to be sure, I never heard an accusation worse than “he gossips a lot” and to my way of judging the attribution even of this mild failing was gratuitous.

Despite my inherited suspicions I came to have great affection for X, and found him one of the nicer and most decent people I met during my ethnographic stay.² The circumstances under which I came to do so are telling. X had been to my host's *ger* several times in order to invite me for a stay at his. This is a local custom: a person may be honored by being hosted and feted for a few days. The host brings one of his own horses to the home of the guest, and they both ride back to the host's *ger*. This happened to me countless times as people wanted their photographs taken. Every time X came by to suggest this, my hostess made excuses for me or deflected the invitation. But X persisted and let her know that he was killing a goat for me. At this point my hostess insisted that I must go. I myself was not keen to go given that he was supposed to be such a bad person. But my hostess was adamant, and explained that if I didn't go, X would blame her and people would think she was a bad person. In other words, despite her strong dislike for X, and despite his supposedly bad name, any perception that my hostess was disrespecting him would have affected her reputation negatively, or so she thought, and this was intolerable.

In his capacity as guest at other people's *ger*, X was not treated with less respect than was due to a man of his age. *Shimiin arkhi* (a home-distilled spirit from cow and goat milk that the nomads drink) must always be served when the host wishes to honor a guest, and for elderly guests it is *de rigueur*. *Shimiin arkhi* was always brought out when X visited H1. And H1 invited X to the inauguration of H1's younger brother's *ger* (a ceremony which anticipates the younger brother's impending marriage), and was treated with great respect.

So this anecdote suggest that Torguuds are worried so much about being perceived as a ‘bad person’ that they dare not transgress even those with a bad reputation. This statement nicely contains the paradox: if the idea of punishing even those who are considered ‘bad’ makes the potential punisher worry that this

² I think X is disliked not because of moral failings, of which I have no evidence—all to the contrary. Rather, I think he is disliked because X is, quite simply, very *weird*. He has a strange, high-pitched, loud theatrical laugh which has a knack for exploding often and in all sorts of situations where other people see nothing funny, resulting in much awkwardness. In addition, X constantly breaks many other small norms of propriety. One particularly extreme transgression is his habit for slapping himself and other people on various parts of the body when he has a fit of laughter. This is common in other cultures and intensifies the shared experience of humor, but among the generally reserved Torguuds *nobody* else does this.

punishing behavior will give him or her a ‘bad name,’ then violators will not be punished. But if they aren’t, then why should anybody fret so much about getting a bad reputation when this carries no tangible costs?

Town-dwellers seem no more likely than nomads to punish others for social offenses. My friend Baajaa, a town-dweller, is typical. Asked what would be the social cost to a person whom all recognize as a *muu khün*, he replied almost automatically: “no cost whatever.” So I tried with an example:

“Okay,” I said, “Suppose that there is this person whom everybody says is a *muu khün*. You are driving your truck and you pass by this person, who signals to you that he is going your way and needs a ride. Do you stop and give him a ride?”

“If he hasn’t done anything to me personally, then yes, I will. If he has done something to me, then I won’t.”

As Baajaa suggests, if punishment is forthcoming it is only after specific offenses and only from those directly offended. The structure of the experiment is interesting, in this respect. A responder cannot affect future proposer behavior with respect to that dyad specifically, because the game is played once, and anonymously. The responder also knows that the proposer doesn’t know the responder’s identity, so ‘bad’ proposer behavior may be interpreted as an offense against the social order, rather than the responder’s person, as in the case of real-life offenses against third-parties. From this perspective, the ethnographic observations do not predict responder punishment, and there almost wasn’t any.

Proposer behavior also seems to match the ethnographic observations. Proposers fret about their behavior in the experiment even though responder punishment is not forthcoming, and in real life people worry much about behaviors that will bring them a bad reputation, even though no tangible costs appear to follow. One interpretation of this is that the same mechanisms which make proposers worry about their reputations in their ordinary lives also operate in the experimental situation. However, there is one discrepancy: in the experiment, proposers *did* actually think there would be tangible costs to ‘bad’ behavior (i.e. they believed responders would irrationally punish low offers), whereas in my open-ended interviews people routinely explained to me that there were no real costs to those with a bad reputation.

I believe this discrepancy, though it appears rather sharp, is in fact illusory. I believe it results from the fact that when I interview people about everyday social punishments, they are asked to reason about hypothetical third-parties, rather than their own behavior. In the experiment, on the other hand, they are reasoning about the consequences of their own imminent behavior, and there may well be a psychological bias to exaggerate the possible negative consequences of one’s own imminent and immoral behavior beyond what the empirical data of one’s experience supports. Why? Perhaps because such a mechanism adaptively keeps people from the costs of social reactions on the tail of a distribution of reactions, which tail they typically will not have seen (but

should avoid). When thinking of *other* people, however, the problem is not strategic but empirical, and the bias may not operate. This hypothesis explains why people are worried about their own reputations even though they judge the bad reputations of others to carry no costs, and it also explains why responders did not share the erroneous intuitions of proposers: responders were not thinking of their own impending behavior, just reacting to somebody else's.

A wrinkle remains: why did proposers say that *they themselves* would have punished low offers? It is quite possible that this is a post-facto rationalization of the trepidation they feel when making the offer. Having told me explicitly that they think responders will punish certain amounts in order to explain their trepidation in making an offer, they hypothesize their own behavior as responders in such a way as to keep their rationalizations consistent. There is evidence that much explicit moral reasoning is a post-facto effort to render one's behavior plausible, rather than causally determinative of moral behavior (Haidt 1993, 2001), and this could be a case in point. This hypothesis has the additional virtue that it explains why the intuitions of proposers and responders were different: responders, by and large, did not think there was such a thing as a punishable offer.

Although the above may help establish a continuity between the ethnographic and experimental data, and therefore add plausibility to the hypothesis that participants frame the experimental situation with the same paradox that rules their behavior in the world, it still does not explain the paradox. We may state the problem as two questions: First, why is there so little social punishment against third parties perceived to be 'bad'?; and second, why do my informants worry so much about their moral reputations when tangible punishment is so seldom forthcoming? I'll tackle each question in turn.

2. Why is there so little social punishment against third parties perceived to be 'bad'?

It is possible that the ecology of nomadic pastoralism explains the paucity of social punishment. Within reasonable costs, punishment in the context of reciprocal interactions can give a clear benefit to the punisher in that it socializes the 'cheater' not to offend the punisher, motivating improved behavior in that particular dyad. But third-party punishment (i.e. for generalized offenses against the social order, rather than against one's individual person) is a public good, and for this reason is susceptible to free-rider effects. However, third-party punishment may give a net benefit to the punisher when it can be witnessed by other members of society, who are the beneficiaries of one's punishment, and who therefore can attest to one's moral rectitude as evidenced by one's exertions to enforce the moral order. In other words, by punishing those who offend third parties, the punisher emits a signal that she is morally upright and therefore a good partner for risky reciprocal interactions subject to cheating temptations. If others then prefer this punisher in all sorts of reciprocal games, such benefits may overcompensate the costs incurred in punishing.

If this is the logic of third-party punishment, then one would expect that as the signal strength of punishment weakens, one should see less punishment, because when other people are not on hand to witness one's acts of social punishment, severing or imperiling one's reciprocal ties to a person may be too expensive to perform if that person has not offended against oneself. In such an environment, a norm holding social punishment *itself* as against the moral order may be favored.

This interpretation receives support from the fact that the behavior of Kazakhs and Torguuds, whether proposers or responders, is statistically indistinguishable. Although Kazakhs and Torguuds are different ethnies with many important contrasts in their interactional norms, they should be very similar in this domain of behavior if it is indeed strongly determined, as hypothesized, by the ecological adaptation. And Torguuds and Kazakhs earn their living in almost exactly the same manner, with a very dispersed lifestyle consisting of relatively isolated family units (Mongolia has the lowest density of any country in the world).

Also supporting this ecological hypothesis is the fact that the other herder groups in the sample, Orma and Sangu, with comparably sized offers, had essentially identical rejection rates, respectively, 0.04 and 0.05; Torguuds had a rejection rate (for the *real* offers) of 0.05, and Kazakhs of 0.10. The Achuar, Mapuche, Tsimané, Quichua, and Machiguenga, although they are not pastoralists, share with the pastoralist groups a relatively low-density arrangement with individual households living in considerable isolation from each other. The Quichua and Machiguenga have mean offers that are too low for the comparison of rejection rates with the herders to be meaningful. But this comparison can be made with the other three, and these rejection rates are 0.14 (Achuar), 0.07 (Mapuche), and 0.0 (Tsimané). By contrast, groups with denser lifestyles and comparable mean offers showed higher rejection rates. The difference between Sangu farmers (0.25) and herders (0.05), is significant, because what distinguishes them here is the spatial arrangement, not so much their cosmology, which they largely share, being members of the same ethnie (see Henrich *et al* this volume, for the complete data set). By contrast, as we saw, Kazakhs and Torguuds, who are members of different ethnies, but have an almost identical spatial arrangement, have very similar (and predictably low) rejection rates.

Although it is true that Torguud town-dwellers have a denser lifestyle (but not really the Kazakhs who, even as 'town'-dwellers, live very dispersed), they did not punish. It is true that the norms in town and among nomads (though differently nuanced) are largely the same because many factors contribute to the bi-directional flow of norms. First, nomads and town-dwellers are each other's kin, and there are many reciprocal ties between nomads and town-dwellers. Second, migration in livelihood goes both ways: many children of nomads become townsfolk and some townsfolk take up pastoralism. Third, even the children of town-dwellers spend the summers in the highlands assisting their herder relatives in order to escape the mosquitoes in the floodplains. Fourth,

nomads are considered experts in tradition so that even though there is a tendency for the townsfolk to relax some norms, there is nevertheless a gravitational pull towards the nomads' standards, which are considered more genuine. But these arguments are not entirely persuasive because Sangu herders and farmers have different rejection rates even though they are members of the same ethnies. I believe the best answer is that the signal value of punishment is not much better in town than it is among the nomads. Interaction in town is far from intense. People mostly keep to themselves and immediate neighbors do not really share much community life unless they are relatives. Town life in fact feels oppressively desolate and ghost-townish, and the town itself is quite spread out, with considerable spacing between dwellings in some of its sections.

The signaling interpretation of the logic of third-party punishment I am defending here is similar to Smith & Bliege-Bird's (2000) account of the provision of public goods following from the benefits of signaling one's quality. In their explanation, public goods are produced if they require qualities that are costly to fake, and if advertising those qualities brings net benefits to the producer of public goods. Moralistic punishment is a public good, but it advertises nothing if it is not observed, which requires a relatively dense way of life. What I have suggested above is that with a sufficiently dense lifestyle, such punishment advertises 'commitment to the moral order,' and this results in benefits to the punisher because s/he will then be preferred in reciprocal interactions due to the lower interaction risk this person is understood to represent. To make this work, however, moralistic punishment should be an *index* for the moral person, and a costly signal for would-be fakers (i.e. hypocrisy should be difficult). The solution to this requirement is at hand if it is the case that abiding by the moral order is a matter of emotional 'commitment' (see Hirschleifer 1987; Frank 1988). That is, if it is one's deep, non-negotiable, and non-calculating emotional commitment to the social norms that *causes* one to punish violations of it directed against third parties, then one expects that the same emotional commitment makes violations by the same punisher unlikely. A 'faker' who lacks the emotional commitment, however, will find it difficult to incur the costs of punishment given that this requires an irrational short-term decision to incur considerable costs to provide a public good.

This signaling/ecological hypothesis may help explain the absence of any—even mild—punishment by third-parties in Bulgan Sum. But a puzzle remains: why should there be a bias to exaggerate the costs to one's own bad behavior—as suggested above—even when one lives in a low-density herding society where 'bad people' are not punished?

Here is one possible answer. Given entry into adaptive zones where the signal value of punishment is low (e.g. low-density nomadic herding), then, as suggested earlier, a norm that 'punishing is bad' will be favored and become part of the moral order, because those who punished when nobody was around to see it would quickly learn that the costs of this behavior were larger than its benefits. But if entry into this kind of adaptive zone is recent and populations that occupy it are less common than more dense societies, and if they are not reproductively

isolated from those in which the signal value of punishment is high, then evolution by natural selection has not been able to remove an innate tendency to develop an emotional commitment to the moral that was selected for in denser social arrangements that are much older. Khazanov's (1994) well-founded argument that nomadic herding is very recent because it depends on nearby agricultural societies for its success lends support to this idea.

In conclusion, if the proximate psychological mechanisms that provide individuals with an emotional commitment to the moral order include a fear for one's reputation (because in the environments of history punishment for 'bad people' has been common, and a fear for one's reputation adaptively prevents such punishment), this may lead to rationalizations about the costs to *one's own* 'bad' behavior that exaggerate the perceived costs beyond what is typically observed when *other* individuals are 'bad.'

My explanations are not exactly elegant, but they are not wholly implausible. Human historical processes, both genetic and cultural (to say nothing of their dual-inheritance interactions; Boyd & Richerson 1985) will produce eddies and whorls; elegance is not always what one sees in the data, so perhaps the best explanations will not always be terribly elegant either.

3. Why no ingroup favoritism?

Proposers made larger offers to *outgroup* ethnics, and responders were not more likely to punish the fake, low, putative 'outgroup' offers (or, for that matter, the real offers). These results are in conflict with the expectations of Social Identity Theory (SIT) and Realistic Conflict Theory, so I revisit the relevant arguments in the context of this result.

SIT (and its more recent offspring, self-categorization theory) posit that there is a generalized bias to favor one's ingroup, regardless of how the ingroup is construed, because one's self-esteem is partly a result of a positive evaluation of those groups in which one is a member (Tajfel & Turner 1979; Turner *et. al.* 1987). This theory began as an attempt to explain a whole collection of results obtained with so-called 'minimal groups.' Such groups are formed by arbitrary criteria (e.g. by flipping a coin) and, at the extreme, they have no ecological validity (even to the point of enforcing complete anonymity between members). It was found that even such 'minimal' groups can promote an ingroup favoritism bias in evaluations of, and in the allocation of positive rewards to, ingroup and outgroup members (the original and classic study is Tajfel 1970; see Tajfel 1982, and Diehl 1990 for reviews). SIT claims that self-esteem rises upon performing discriminatory behaviors with an ingroup favoritism bias, and the preferential allocation of benefits to ingroup members is performed in order to achieve that rise in self-esteem.

The problems with this research are coextensive with its strengths. Establishing that discriminatory behavior can take place between any kind of social categories seems important, but any real significance is contingent on allowing for broader interpretations. If what motivates the behavioral bias with

minimal groups in the lab is very weak (as suggested by the very weak effect typically obtained), then perhaps it will get swamped by other biases in the context of more realistic group contrasts and situations. In such a case, we have no more than an interesting exotic result, relevant only to the exotic experimental context of minimal groups. Thus, the purported importance of the minimal group literature rests on the claim that here is evidence for a general ingroup favoritism bias applicable to any and all intergroup contrasts, and therefore helpful for understanding all sorts of intergroup conflict, such as ethnic conflict. Such claims have indeed been made (Tajfel & Turner 1979). But such a powerful 'ingroup favoritism' bias predicts that in the UG there will be a bias to favor the ingroup responder, who will get higher offers than the outgroup responder. This prediction is not incompatible with proposers who wish to make money, it merely says that proposers have a bias making them more reluctant to profit at the expense of the ingroup rather than the outgroup responder.

Consider first that the UG with an ethnicity manipulation forces the proposer to explicitly consider costs and benefits to self (the larger the share for the responder, the smaller the remainder is for the proposer), something that is absent in the minimal group experiments. That this might be important is suggested by the classical minimal group experiments themselves. The ingroup favoritism differential is typically quite small, suggesting that perhaps a larger difference would make the allocator feel bad for having been too unfair to the minimal outgroup member. If ingroup enhancement were the only consideration, then surely participants would be choosing to maximize the difference in favor of the ingroup. Dramatic evidence that allocators indeed weigh costs and benefits to their sense of 'fairness' comes from a study by Mummendey et al. (1992), where participants were found to be scrupulously fair in their allocation of a high-pitched blast of noise, which happens to be a sizable negative reward when compared to the symbolic points or almost worthless rewards typically used in minimal-group experiments. Thus, if people want to be 'fair,' a consideration of sizeable costs and benefits, which the design of the present experiment forces, may be partly responsible for erasing an abstract 'ingroup bias,' if indeed such a thing exists.

But the 'ingroup bias' was not here merely erased. The effect goes in the *opposite* direction, and this requires explanation. The apparent mystery is compounded by the fact that these are ecologically valid groups, drenched in reality in a way that the minimal group paradigm explicitly tries to avoid. Intuitively, one would expect an ingroup bias detectable in the context of minimal groups to grow as reality is added to the group contrast, but the opposite happened.

Realistic Conflict Theory (RCT) explicitly tries to explain the behavior of realistic groups when they find themselves in conflict. This theory "argues that groups become prejudiced toward one another because they are in conflict over real, tangible, material resources" (Sabini 1995:104). Such prejudice in turn motivates discrimination with ingroup favoritism for both positive and negative rewards. The *locus classicus* of RCT is the famous Robber's Cave experiment

(Sherif et al. 1961) where it was found that pitting two teams of adolescents (the teams were formed at random) in different zero-sum contests increased hostility between them, and led to ingroup cooperation for inter-group competition. The converse was also shown: giving them a common goal that could only be solved through cooperation reduced hostility and even caused friendly intergroup relations toward the end. From the perspective of this theory, I should have found ingroup favoritism greater than the meager differentials found in minimal group experiments. After all, a resource is explicitly at stake.

There are two important differences between the groups in the Robber's Cave experiment, and the groups in the present Ultimatum Game: (1) In Robber's Cave the groups are formed at random and have a shallow history; in the present study the groups coming into the experiment are anything but random and they have a very long history; (2) In Robber's Cave realistic conflict was experimentally manipulated, and the conflict, created as it was by the rewards at stake, was immediately palpable and structurally clear; in the present experiment, realistic conflict must be inferred from the ethnographic situation and is less directly palpable and structurally clear. The first point makes Robber's Cave a less good case for realistic conflict than the present study, and the second point a better one.

Do these considerations cancel out exactly? It is impossible to say. What *can* be said is that the present study has groups for which a *sense* of realistic conflict is plausible. The district of Bulgan Sum is composed primarily of Torguud Mongols and Kazakhs. It would be remarkable if Torguud Mongols, living as they are in a state calling itself Mongolia, with an overwhelming majority of ethnic Mongols, did not have at least a vague proprietary feeling about the territory—especially considering that right next door is a state calling itself Kazakhstan where ethnic Kazakhs—at least until recently—could automatically get a passport if they so wished. Moreover, Kazakhs are expanding faster than Mongols because they reproduce more than twice as fast and migrate away only about a third as much as Torguuds do, or less. Again, it would be remarkable if Mongols didn't at least vaguely perceive this as a form of 'invasion'. Where the Mongols are 'proprietary', the Kazakhs could be expected to be 'defensive' about the very same issues. Moreover, in the 1940's there was much raiding and killing going on by bands of outlaws that were eventually driven back behind the Chinese border. The outlaws were of Kazakh ethnicity (followers of a charismatic bandit by the name of Osman). Some of my ethnographic observations suggest that this historical episode does not aid interethnic trust on the part of the Mongols, and neither does the fact that the land across the neighboring Chinese border is populated by Kazakhs. It is certainly true that interethnic relations are on the surface quite friendly, if for the most part distant, and it is not difficult to find positive stereotypes of the Kazakh among the Mongols (as hard-working, responsible, etc.; of course, these could have a double-edge, because it is precisely by doing well that the Kazakhs will take over the territory). Positive stereotypes of Mongols are much harder to find among the Kazakh. What I am arguing for is no more than a vague sense by local actors that the groups are in conflict, even if in a weak and diffuse sense.

From the point of view of RCT, this *could* lead to an outgroup favoritism bias in proposers. If the perception of conflict makes responders more likely to reject outgroup offers, *and proposers anticipate this*, then they might make higher offers to the outgroup responder. However, to do so would not be an ingroup favoritism strategy for the proposers. So this explanation has the proposers caring more about making money than favoring the ingroup, but not the responders. It is not a straightforward prediction from RCT, and neither does it seem plausible. Besides, the behavior of responders showed no such bias so the anticipation of proposers would have to be imagined rather than empirically based (though, admittedly, this is far from impossible, in principle).

The way to maximize satisfaction of an ingroup favoritism bias, in the absence of competing biases, would be for proposers to try to offend outgroup responders but make fair offers to the ingroup responders. Every proposer should reason that low offers to the outgroup and fair or hyperfair offers to the ingroup will result in either (1) low gains for the outgroup with large gains for the ingroup, or (2) zero gain for the outgroup with moderate gains for the ingroup. Responders should reason that any offer from an ingroup proposer should be taken but that any offer from an outgroup proposer should be rejected. Adding a preference for money will weaken the size of these effects but should not tilt them in the opposite direction. However, the observed effects are in the opposite direction.

Thus, neither theory just considered can easily account for the fact that outgroup responders get larger offers. This should be troubling to SCT and RCT given that the present experiment has ecologically valid groups and presents players with cost/benefit decisions more like those they encounter in the real world. The result obtained here suggests that the experimental evidence on which SCT relies is indeed exotic, and that more work is needed to fill out RCT.

What can account for the outgroup bias found in my experiment? One possible answer is that ethnies are norm-groups (see Barth 1969) and in human cognition they are *naturalized* norm-groups, where the differences in normative behavior are believed to follow from what is imagined to be an inherent nature (Gil-White 1999, 2001a, 2001b, 2002a, 2002b). Thus, a Torguud proposer, say, may reason that she has a pretty good guess of how low she can go before a Torguud responder punishes, and that an offer only slightly above that can be made with little risk of rejection. However, the punishing threshold of a Kazakh responder—underlain as it is by a foreign ‘nature’—is something they feel less equipped to guess. These considerations, plus the reasonable assumption that participants follow a return-maximizing strategy within the framework of assumptions they make about the world, may lead proposers to ‘play it safe.’ If they have no way to judge where the outgroup punishing threshold is, they may take what they believe is the ingroup punishing threshold as an anchor, and either offer at that level, or a little higher.

There is no opportunity for this sort of calculation in minimal group experiments because (1) there is no conflict between favoring oneself vs. the other player (by contrast, in this UG the opportunity cost of a bad strategy was

large because the money at play was about a week's salary); and (2) group norms are irrelevant. In the Robber's Cave experiment, a conflict between individual and group benefit does exist, but in point of fact there was very little free-riding there. The difference in the Ultimatum Game may be that the game is played anonymously, and so the lack of monitoring strengthens selfish motives. This consideration suggests that RCT and the hypothesis just offered are not necessarily in conflict. It may be simultaneously true that ethnic groups are perceived as naturalized norm groups, and that realistic conflict will promote discriminatory ingroup altruism when there is good monitoring.

Not every participant was articulate under cross-examination about why the outgroup had received a better offer. But from those who were, I obtained two main categories of responses, which I paraphrase and stylize: (1) "I know my coethnics. I know how low I can go without getting a rejection. Them [outgroup members] I don't really know, so I play it safe"; (2) "I don't want to foster any misunderstanding between the groups." Another plausible reason for giving higher offers to the outgroup is an anticipation that outgroup members will be likely to punish low offers *coming from an outgroup member*. However, nobody gave such an explanation. The two main categories of introspections are not necessarily mutually exclusive. Take, for example, the player who made the only *hyper-fair* offer in Study 2. This was a very polite and agreeable Kazakh man, and his hyper-fair offer went to the outgroup responder (0.56). After I suggested that he had made such a 'nice' offer because he wanted to make sure the responder accepted and he got his money, he replied: "That, in the first place. But, in the second place, I don't want to hurt people's feelings." I do not doubt his sincerity, but his co-ethnics have feelings too, and he only offered 0.38 to the ingroup responder. The fact that 'people's feelings' are more carefully respected when they belong to an outgroup member speaks volumes about the plausible instrumentality of such considerations: players want, first and foremost, to make money, but they err on the side of safety when they feel insecure about the location of the punishing threshold. If ethnic groups are norm groups, and our psychology has evolved to deal with this (Gil-White 1999, 2001a, 2002b) then a feeling of insecurity about the location of the punishing threshold is more likely with the outgroup responder.

This, of course, does not mean that certain ingroups are not favored discriminatorily in the real world, nor does it mean that ethnies are not among the social categories that more easily generate such ingroup favoritism. Evidence for both propositions is already overwhelming from the historical and current data on nationalism, etc. What this result suggests, however, is that an abstract 'ingroup favoritism bias,' applicable to any and all ingroups is weak at best, and not particularly illuminating as the basis of an account of intergroup discrimination and conflict.

IV. Conclusion

The results of the ultimatum experiments in Bulgan Sum support the hypothesis that performance in this game is significantly influenced by local culture and the ideas concerning fair offers and punishment contained therein. Proposers were fearful of possible responder punishment, but a taste for punishment failed to manifest itself among the responders. This is counterintuitive in light of the results which have been obtained in these experiments when conducted with Western (or Westernized) populations. However, they are plausible and comprehensible when evaluated against the backdrop of a local culture that stresses the importance of one's reputation as a moral person, but considers active punishment itself immoral.

Two major theories in social psychology—Social Categorization Theory (SCT) and Realistic Conflict Theory (RCT)—have tried to explain intergroup animosity with experimental setups that involved groups very different, and decision matrices very alien, to the ones that ultimately one wishes to understand (e.g. actors who weigh costs to themselves in the context of interethnic conflict). The results of this experiment suggest SCT has not given us a very good foothold into the cognitive processes important to intergroup conflict, and that RCT is incomplete and more attention should be given to the role of monitoring.

In the future, economic experiments will probably offer a better avenue to an investigation of competing claims about the cognitive basis of intergroup conflict. If a good theoretical model must be a good caricature of a real causal process, then a good experiment must be a good caricature of a real situation (Gil-White 2001c). The paradigm of experimental economics has the advantage of allowing for decision matrices involving realistic costs and benefits. So far, however, the experimental literature in economics has mostly neglected the possible effects of social category membership on decisions. Given the fact that many decisions that we make are framed in terms of the roles and categories we inhabit as social actors, this is a question obviously deserving of close scrutiny. The present study produces no definitive answers, but it suggests that here lies a rich arena for further work.

References

- Barth, F., ed. 1969. *Ethnic groups and boundaries: The social organization of cultural differences*. Boston: Little Brown & Co.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology*, 1, 263-92.
- Gil-White, F. J. (1999). How thick is blood? The plot thickens...: If ethnic actors are primordialists, what remains of the circumstantialist/primordialist controversy? *Ethnic and Racial Studies*, 22, 789-820.
- Gil-White, F. J. (2001a). Are ethnic groups biological 'species' to the human brain?: Essentialism in our cognition of some social categories. *Current Anthropology*, 42, 515-554.
- Gil-White, F. J. (2001b). Sorting is not categorization: A critique of the claim that Brazilians have fuzzy racial categories. *Journal of cognition and culture*, 1, 219-249.
- Gil-White, F. J. (2001c). A good experiment of choice behavior is a good caricature of a real situation. *Brain and behavioral sciences*, 24, 409.
- Gil-White, F. J. (2002a). The cognition of ethnicity: Native category systems under the field-experimental microscope. *Field Methods*, 14, 170-198.
- Gil-White, F. J. (2002b). Is ethnocentrism adaptive?: An ethnographic analysis. under review, *Nomadic peoples*.
- Frank, R. H. (1988). *Passions within reason: the strategic role of the emotions*. New York: W.W. Norton.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *In press, Psychological Review*.
- Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality & Social Psychology*, 65, 613-628.
- Hirshleifer, J. (1987). On the emotions as guarantors of threats. In J. Dupré (Ed.), *The latest on the best: Essays in evolution and optimality*. Cambridge MA: MIT Press.
- Khazanov, A. (1994). *Nomads and the outside world*. Madison, WI: University of Wisconsin Press.
- Mummendey, A., Simon, B., Dietze, C., Grünert, M., Haeger, G., et al. (1992). Categorization is not enough: Intergroup discrimination in negative outcome allocation. *Journal of experimental social psychology*, 28, 155-44
- Sabini, J. (1995). *Social Psychology*. New York and London: Norton.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robber's Cave experiment*. Norman: University of Oklahoma Book Exchange.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96-102

- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual review of psychology*, 33, 1-39
- Tajfel, H. & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.). *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford and New York: Blackwell.
- Smith, E. A., & Bliege Bird, R. L. (2000). Turtle hunting and tombstone opening: Public generosity as costly signaling. *Evolution and human behavior*, 21, 245-261.
- Szynkiewicz, S. (1993). Mongolia's nomads build a new society again: social structures and obligations on the eve of the private economy. *Nomadic peoples*, 33, 73-103

Appendix A: Protocol for proposer recruitment in Study 1

Note, I refer to myself (Fran, my nickname) in the third-person throughout because of the role-playing used in the explanations. In the explanations, I pretend to be one of the participants, and Boldoo, my assistant, the other, for illustrative/explanatory purposes.

“Fran is doing an experiment. This experiment is like a game. I will explain shortly. If you participate, I will pay you T1600.”

- 1) “This experiment has been done many times, but always in America or else in Europe, and always with city people. For this reason some anthropologists are going to Asia and Africa and will do this experiment again with country folk.”
- 2) “Okay. Boldoo and I will now show you how to do this. This experiment is like a game, and you can make some money.”
 - “For example, make believe I am not Fran. Okay? Right now I am this guy Tsatsral, a Torguud. Tsatsral and Boldoo will play together now.”
 - “Fran gave Tsatsral these T8000. Tsatsral is the proposer, and Boldoo is the responder. Since Tsatsral is the proposer, with this money I will make a partition. I will keep some of the money, and the rest I will offer to Boldoo.”
 - “Since Boldoo is the responder he will get to decide once I make my partition. If Boldoo allows my partition, he will get the money I offered, and I will get the money that I kept. But if Boldoo does not allow my partition, then Boldoo and Tsatsral get nothing, and Fran will again collect this money.”
- 3) [Here I bring the money out and I do the example visually by placing the offer at Boldoo’s feet, and my offer at mine]
 - “For example: After getting T8000 from Fran, I will partition the money thus: I keep T6500, and to Boldoo I offer T1500.”
 - “Boldoo is the responder, so if he allows this partition, he will get those T1500, and I, Tsatsral, will get these T6500. But if Boldoo does not allow this partition, nobody gets anything and Fran will again take all this money back from us.”

4) “All right. Now Boldoo and I will play three examples.” [The examples are all acted out. I pretend to take the money from an invisible Fran to my left and explain that me, Tsatsral, is taking the money from Fran. Then I proceed to make the partition and, once done, I ask Boldoo whether he will allow it. The partitions and Boldoo’s scripted answers are as follows].

- Offer: T1500. Response: Won’t allow.
- Offer: T4000. Response: Will allow.
- Offer: T3000. Response: Will allow.

[After each response, I explain exactly what that means. In the first example, after Boldoo says he doesn’t allow it, I collect all of the money and pretend to give it back to the ‘air’ Fran to my left, explaining that since the offer was not allowed, nobody gets any of this money and it all goes back to Fran. In the second example I explain that since Boldoo allowed the partition, he gets exactly T4000, the money offered, and I get exactly T4000, making a motion for Boldoo to pocket the money and doing the same with my money myself. The third example is as the second except for the quantities apportioned.]

5) “Now we will ask *you*.” [This is the test. I present them with the following examples and ask them to tell me what will happen. Every example is acted out in full just like the previous ones.]

- “If Tsatsral makes the following partition: T6000 for Tsatsral, and T2000 for Boldoo, what will happen if Boldoo allows the partition? (...) What will happen if Boldoo doesn’t allow the partition? (...)”
- [As above with T4500 for Tsatsral, and T2000 for Boldoo].

6) “Okay. You’ve understood now. But this is not exactly the real experiment. We will now explain the real experiment.” [At this point I lay on the floor the twenty pictures of the responders.]

- “You will be a proposer. Among these twenty people is your responder. I cannot tell you which one is your responder because you may not know that. Everybody will play anonymously.”
- “You will now make a partition. After you make the partition, I will go to the responder’s home and will show him/her your partition. [This was explained with role-playing again where I pretended to go to Boldoo’s *ger* (who is now standing in for the responder) and went through the motions of how the partition

would be explained to the responder.] Then, that person will decide: he/she will either allow your partition, or not.”

- “Your responder may not know who *you* are. I will show the responder twenty pictures, including your own, but I will not tell him/her who the proposer is. All participants will always be anonymous.”
- “Other than who you are, the responder will know everything. He/she will know how many *tögrög* I gave you. I will also tell him, of those T8000, how many you offered and how many you propose to keep. I will also explain to the responder exactly what will happen if he/she allows, or doesn’t allow the partition. The responder will know everything. The only thing the responder doesn’t know is who you are.”

Appendix B: Protocol for responder recruitment in Study 2

Start with the protocol for responder recruitment used in Study 1. Once finished, administer the test shown below. Every one of the smallest-level points represents an item of information that is important for the responder to understand. Every answer following each of these questions is the information that the responder should supply, and the explanation that the tester must give should the responder's answer fail. Only moderately-good-to-excellent performance on the first pass is acceptable. Anybody failing badly on the first pass, or showing great difficulty in understanding the points she got wrong, must be excluded from participation in the experiment. Those who fail on the basic points concerning the mechanics of the game should be considered to have failed miserably, and should not be allowed into the experiment. (Note: The test for proposers was essentially identical, with obvious modifications to make it specific to proposers.)

- I. Tests whether responders understand that the money is not the proposer's, or mine, but the school's, and that the proposer is offering money because otherwise she may not play.
 - **Why will the proposer offer you money?** A: The answer to this question could be anything at all, and will guide which of the following questions is used, or which is used first. If there is long hesitation just jump into the following specific questions.
 - A. **What happens if the proposer doesn't want to make a partition?** A: He may not play unless he makes a partition. If he doesn't want to make a partition, he may not play.
 - B. **Whose money is the proposer playing with?** A: He got money from Fran to play with. This is not Fran's money. The school gave Fran money to do this experiment with.
 - C. **Are we playing with real or pretend money?** A: Real money.
- II. Tests whether responders understand the mechanics of the game and the consequences of their decision.
 - **What will you do after I show you the proposer's partition?** A: The answer to this question could be anything at all, and will guide which of the following questions is used, or which is used first. If there is long hesitation just jump into the following specific questions.
 - A. **After seeing the proposer's partition, may you choose not to allow it?** A: Yes.

- B. **After seeing the proposer's partition, may you choose to allow it?** A: Yes. Whether I allow or not the partition is my decision.
- C. **If you don't allow the proposer's partition, what will Fran do?** A: Fran will not give me any money, but he will not give any money to the proposer either.
- ⇒ **Where does the money go, then?** A: Fran will give it back to the school.
- ⇒ **If you don't allow the partition, does the proposer make another offer?** A: No. We play only once. The proposer only offers once, and after I make my decision that is the end of the game. The proposer cannot make another offer.
- D. **If you allow the partition, what will Fran do?** A: Fran will give me the money offered to me by the proposer, and he will also go back to the proposer's *ger* and give him the remaining money.
- ⇒ **Does the school care that it may never see this money again?** A: The school does not care at all. I must not decide out of concern for the school. The school gave this money to do the experiment and it doesn't need the money back.
- E. **How much money will Fran give the proposer to play with?** A: T8000.
- F. **Does the proposer keep that money?** A: No. Fran gives it to the proposer merely to see the proposer's partition, then Fran takes the money again. The proposer gets nothing unless I allow the partition.
- G. **Can the proposer make any partition that she wants?** A: Yes.
- H. **Will you and the proposer play with real or pretend money?** A: For real money.

III. Tests whether the responder understands what the proposer knows.

- **What will I tell the proposer?** A: The answer to this question could be anything at all, and will guide which of the following questions is used, or which is used first. If there is long hesitation just jump into the following specific questions.

- A. **Does the proposer know who you are?** A: No.
Everybody is playing anonymously.
- B. **Does the proposer know what happens if you allow the offer?** A: Yes. The proposer knows that if I allow the offer, Fran will give me the money offered, and he will also give the remainder to the proposer.
- C. **Does the proposer know what happens if you don't allow the offer?** A: Yes. The proposer knows that if I don't allow the offer, Fran will not give me any money, but he will not give any money to the proposer either, and will return the money to the school.
- D. **Will Fran explain to the proposer what the responder knows?** A: Yes. Fran will tell the proposer, "The responder knows I gave you T8000 to play with." Fran will also tell the proposer, "The responder knows what happens if she allows or does not allow your offer."
- E. **What is the only thing that the proposer doesn't know?**
A: My name.
- F. **What is the only thing that you don't know?** A: The proposer's name.