

Gil-White, F. J. 2004 (In press). Ultimatum Game with ethnicity manipulation: Problems faced doing field economic experiments, and their solutions. *Field Methods*.

Ultimatum Game with ethnicity manipulation

Problems faced doing field economic experiments, and their solutions

Francisco J. Gil-White

Psychology, University of Pennsylvania
3815 Walnut Street
Philadelphia PA 19104-6196
fgil@psych.upenn.edu
<http://www.psych.upenn.edu/~fgil/>

Introduction

This paper reports two Ultimatum Game experiments (the first, a pilot, the second, a successful experiment) conducted with Mongols and Kazakhs in Western Mongolia, with special attention to the methodological difficulties that pertain to economic experiments in the field.

One purpose of the study was to test how non-Westerners, living in a simple, non-industrialized society, would perform in an Ultimatum Game (UG), and particularly if they would do so differently when paired against coethnics than when paired against outgroup ethnics.

In the UG, a “proposer” is paired against a “responder.” The first receives a sum of money from the experimenter with which to make an offer to the responder, proposing to keep the rest for himself. The responder inspects the offer and then decides to make either one of two responses: (1) “Accept,” in which case the money is divided as specified in the proposer’s offer; or (2) “Reject,” in which case neither player gets any money as it all goes back to the experimenter. The players are usually out of sight of one another, with information exchanged by computer. They play anonymously.

It is a robust finding in the industrialized world, and especially in Western settings, that offers below about 20% of the total available tend to be rejected, even when the absolute amount at stake is large. Proposers appear to anticipate this and as a result “offers typically average about 30-40 percent of the total, with a 50-50 split often the mode” (Camerer & Thaler 1995). This result contradicts the expectations of neoclassical economics, in which rational, self-interested responders, with a utility for money, should accept any amount of free money,

and where rational, self-interested proposers – again, with a utility for money – should anticipate this and offer the lowest possible amount.

In a recent study by Henrich (2000), Matsigenka Indians of Peru, in both the proposer and responder roles, performed – unlike those in industrialized settings – relatively close to neoclassical expectations, suggesting that perhaps the corpus of results obtained so far reflect the norms of industrialized market societies, and perhaps Western societies in particular. Similar to the Matsigenka, rural Mongol and Kazakh responders were reluctant to punish low offers. And yet proposers were nevertheless very careful not to offend them, and made offers well-above the empirically ascertained income-maximizing offer. This suggests that testing other non-Western, non-industrialized populations may yield results of comparative interest (see Henrich et. al. 2003, for results from 16 small-scale societies). In this article, I will lay out the method I used for conducting UG experiments in Mongolia. The method can be adapted for use in other cultures.

A large literature in social psychology has found what is called an ‘ingroup favoritism’ bias – that is, a bias to discriminate against outgroups and in favor of ingroups. The experimental paradigm used is that of ‘minimal-groups.’ In such experiments, That is, people are asked to allocate rewards to ingroup and outgroup members, but the ‘groups’ are categories purposefully stripped of all possible social meaning. For example, subjects might be told that they were assigned to one group or the other on the basis of the flip of a coin. Typically, subjects allocate symbolic points or monetary rewards so small as to be arguably indistinguishable from symbolic points. The original and classic study is Tajfel (1970); see Tajfel (1982), and Diehl (1990) for reviews.

Pertinent questions are the following. First, will the ‘ingroup favoritism bias’ survive decisions made over larger stakes? Second, what happens when people have to *pay* for their allocation decisions? Third, does the ‘ingroup favoritism bias’ increase when social categories are real and least weakly competitive, as in the case of local ethnies?

The results of the experiment reported here contradict the expectations of the Social Identity Theory (SIT; see Tajfel & Turner 1979) or Social Categorization Theory (SCT; see Turner et. al. 1987) which emerged to explain the ‘minimal-group paradigm’ results. Proposers made larger offers to *outgroup*, not ingroup, members, and responders were not more likely to punish outgroup than ingroup members for low offers.

I. Ethno-geography

Bulgan Sum is a district in the province of Hovd, Republic of Mongolia. The field site 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 who roam the district, and it is also their 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 Bangyakhlan clan, with whom I worked, migrate quite far to their summer pastures. In the process, they traverse territory belonging to other clans.

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 Bangyakhlan 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 Bangyakhlan 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 Bangyakhlan 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.

These days, the country folk in Bulgan Sum don’t earn wages and the townsfolk do—that is, those who have a job. In town, unemployment is supposedly rampant but, except for government jobs, most of 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 property insurance and 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 use fewer camels than Torguuds and typically do not breed them. Thus, they must buy them from

Torguuds. The same is true for horses, most of which are bred by Mongols. 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.

The money at stake in the ultimatum game conducted in Bulgan sums to T8000, about \$10 USC at the time, and the equivalent of a few days wages for a hired hand (see Gil-White 2003).

II. 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 methodological lessons learned from it, help increase the confidence for the results obtained in study 2.

This first experiment was conducted exclusively with Torguud nomads from the Bangyakhlan 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 consisted 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 quick-fire 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. (Note: the point of the photographs is to reassure participants that their counterparts exist and have been contacted, and that the whole thing is not a con.)
9. A final trip was made to each proposer's tent. The responder's response was conveyed, and any money due was apportioned.

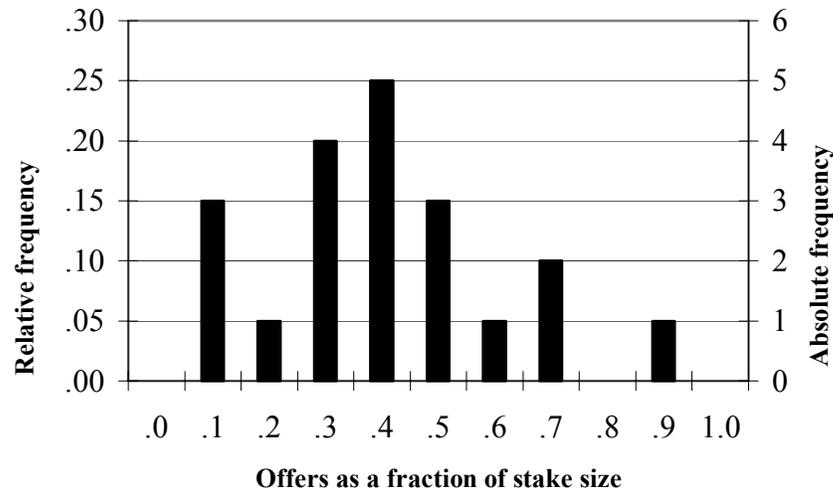
Appendix A has the details of the methodological protocol for the explanations used during proposer recruitment (the explanations to the responder were essentially identical, with a few obvious modifications necessary to make it specific to responders). I was always left alone in the *ger* with the proposer/responder 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 punishing threshold did they anticipate? (2) Was this answer consistent with their reasoning in the experiment, and with their offers? (3) What is the lowest amount that the proposers themselves would have accepted, had they been responders? (4) Of those responders who had rejected the offers made to them, why had they? 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 the distribution in **fig.1** are not too surprising from what we know of experiments carried out in developed settings. 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. There is 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, who explained under a second cross-examination (resulting from Boldoo's amused sharing of his father's confidences with me) that the old man did not feel he could take money from a poor student like myself. (Normally, I would not have recruited the relative of my assistant to be a subject, but the shortage of subjects was quite severe).

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 absurdly 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. But the day after, during informal relaxation with Boldoo in Mönxbat's *ger*, Boldoo let out a big laugh when he understood from Mönxbat that the latter had assumed the game was merely to get the responder to agree to the partition. He thought the money was make-believe, and had been very surprised when I brought him back his T500 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.

My explanation protocol had been careful, step-wise, gently paced, and heavily demonstrative, and 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!). Despite my care, I had failed to anticipate the size of the 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.

Upon cross-examining the responders my skepticism only grew. 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. Moreover, 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).

I hypothesized that perhaps responders had not understood that the proposer *had* to play (that is, they may have imagined that one possible move for the proposer was to make no offer and pocket the money, without the responder getting to play). If so, they might have regarded any offer as a gift, and missed the point of the game. A corollary hypothesis was that perhaps those socialized into the proposer role had an easier time understanding the game, which would

explain the discrepancy between responder behavior and the answers proposers gave as hypothetical responders.

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 consisted 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 were Torguud and half Kazakh. Of the 20 Torguuds, 10 were proposers and 10 responders, and likewise for the Kazakhs. The ratio of males to females per ethnic group per role, in both cases, was 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 had to make, as well as the two decisions that faced every responder. Proposers and responders were assigned at random.

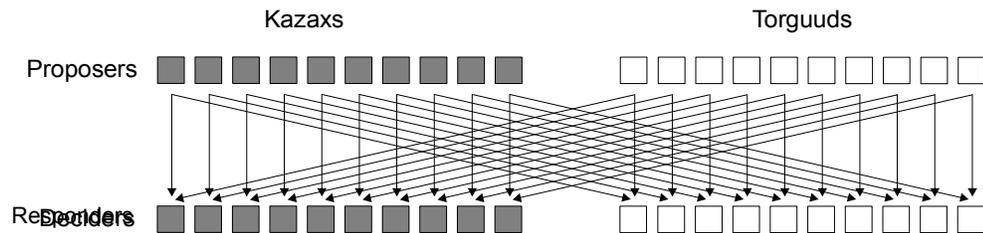


Fig. 2. Structure of offers and responses

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 10 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 10 photographs was presented (half of the participants made an ingroup offer first, and the other half made an outgroup offer first).¹ To counter-balance, half of all proposers, at random, made an offer to an ingroup responder *first*, and the other half to an *outgroup* responder first (order had no effect²).
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 10 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

¹ If I were to do this again, I would include an admonition to make sure the proposer understood, when told about the second game, that the responder in the second game, like the one in the first, would see the photographs of the proposers. It seemed clear that they understood this, but the question is important and an ideal methodology would guarantee it.

² Each reported mean is from a sample with size $n=10$. Offers to *outgroup* responders are as follows: “ingroup first” mean offer to outgroup = 3350, “outgroup first” mean offer to outgroup = 3450; $t = -0.296$, $p = 0.77$. Offers to *ingroup* responders are as follows: “ingroup first” mean offer to ingroup = 2950, “outgroup first” mean offer to ingroup = 2700; $t = 0.8$, $p = 0.44$;

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 the same respect and consideration accorded to Western subjects may seem laudable, but consider the cultural context and the fact that my behavior not only jeopardized the experiment, but unsettled the participants. The lesson is that the obligation of cultural sensitivity to our participants in faraway places can legitimately involve abandoning practices that are considered obligatory back home.

In Study 1 I found little evidence of a punishing threshold, whether in performance or in answers to hypothetical questions. To test whether this had been a problem of comprehension or an attribute of the population, 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 then 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 record it on my chart and 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.
5. Half of all responders, at random, saw the outgroup offer first, whereas the other half saw the ingroup offer first, to counterbalance. Every responder saw only *one* fake offer, and it was always presented as coming from whichever proposer came first.

In this manner I obtained very good data for reactions to the lowest possible offers (naturally, at the cost of pristine data for the real offers, since arguably being faced with an initial lowest-possible offer may affect what they subsequently did with the real offers).

I should note that making up fake offers is considered a cardinal methodological sin in experimental economics because this arguably "pollutes the pool." In other words, the more one deceives participants in experiments, the less naiveté there will be in the pool from which participants are drawn, and this is

costly to the research because deception requires naiveté. Given that ordinary people have grown quite suspicious about what psychologists are *really* up to in experiments, economists care greatly about their methodological reputation as scientists who never use deception to answer experimental questions.

I think there is a place for deception when no practical alternatives exist to obtain the data one needs, but experimental economists have a good argument that this should be where one goes last. In this particular case, I fault myself for not having been sufficiently creative and/or informed, as there are methods available to obtain the data I wanted without deception. For example, I could have used the method where the offer is placed in an envelope and the responder is asked to declare—before seeing the offer—what the ‘punishing’ threshold is. That is, the responder must say offers below which amount will be rejected.

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 understood 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 - on purpose - in my explanatory routine. Another 5 made one or two mistakes, but not on the basics, and had no trouble understanding the right answers when I explained them. I had reservations about the performance of only two proposers. The numbers for responders are virtually identical.

Table 1 shows the means of Torguud offers to an ingroup and an outgroup responder, as well as the corresponding Kazakh means (each sample n=10), with the Epps-Singleton test³ results comparing Torguuds and Kazakhs.

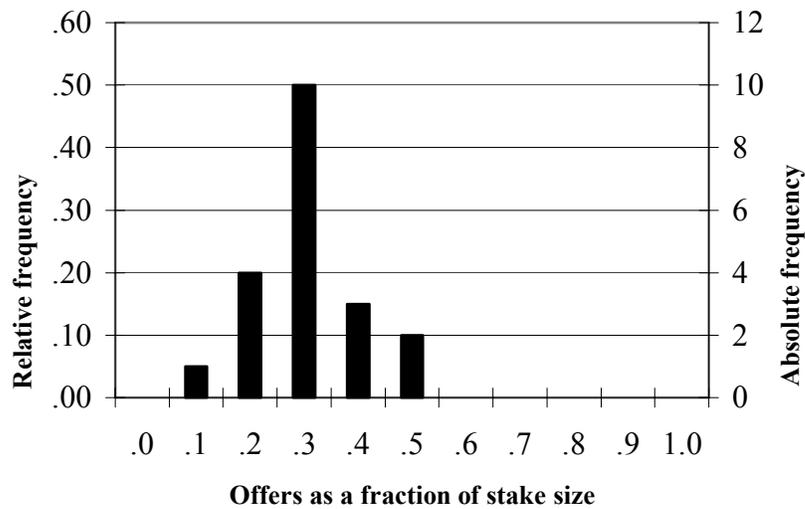
Table 1. Torguud and Kazakh offers in Study 2

	Torguud	Kazakh	CF	p
ingroup offer mean	0.350	0.356	2.63	0.62
outgroup offer mean	0.418	0.431	6.08	0.19

Is the increased confidence in participant comprehension reflected in corresponding changes in the data? First, notice that Mongol and Kazakh distributions of offers are not statistically different, so I combine *all the ingroup offers in Study 2* (see fig.3 for the distribution) in a sample n=20 in order to compare them to the offers in Study 1 - all of which were also ingroup offers (see Table 2 for the comparison).

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

³ Epps & Singleton (1986).



As shown below, the comparison yields a marginally significant difference which, coupled with the ethnographic and interview data, convinces me that a better grasp of the game in Study 2 has indeed produced different results from those in Study 1 (for a visual assessment, compare the distributions shown in **fig.1** and **fig 3.**), and that we are shy of what is usually considered a probability for a chance difference small enough that it can be dismissed ($p=.05$) only because of the small sample size. This is a good example of what I have argued elsewhere is the holistic approach to the interpretation of data in field experiments (Gil-White 2002).

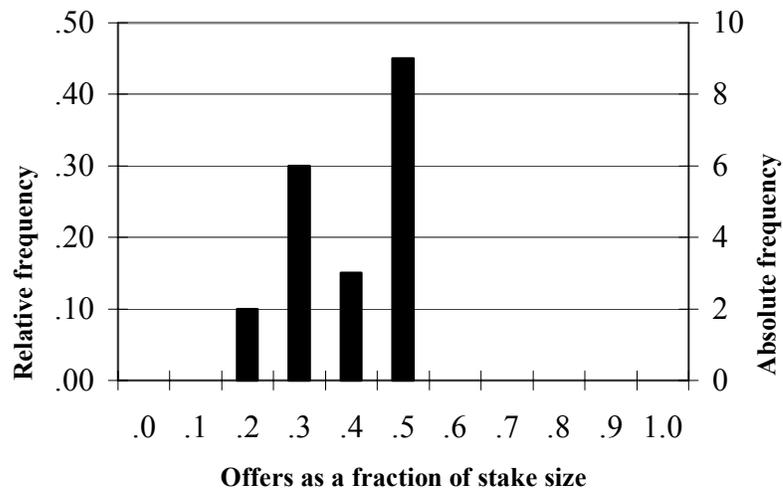
Table 2. Study 1 vs. Study 2* ingroup offers

	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

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.



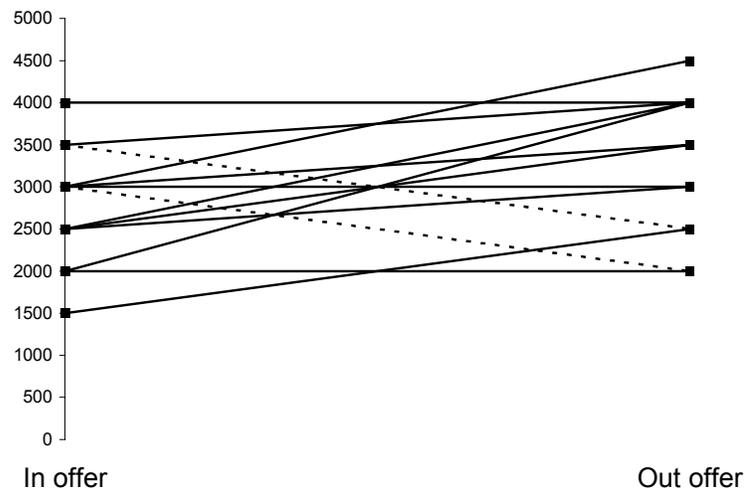
Comparing the distributions in figures 3 and 4 (Table 3) shows them to be different.

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 counterpart, and the other half from the outgroup. Their responses are in Table 4.

Table 4. 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 rejected less often. Table 5 shows the three responders who rejected real offers.

Table 5. The three responders who rejected real offers.

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 (R) an ingroup offer but accepted (A) 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. In summary, the data from rejections yielded no evidence of group favoritism, whether ingroup or outgroup. But the failure again to find any evidence of ingroup favoritism must be noted.

One thing does emerge very clearly from the rejection data. What appeared in Study 1 as a counterintuitive and skewed result, due to methodological error, is in fact reliable: the offers of proposers are considerably higher than what might be expected given the very low punishing thresholds found.

III. General discussion

(For a more in-depth discussion of the results of the experiment, see Gil-White 2003).

It is puzzling that the offers of proposers are so high when the punishing threshold is so low. This could be because people have a stronger preference for avoiding risk than they do for maximizing income. More plausibly, although people wish to maximize income, they may 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.⁴ Immediately after they had made their offers, I queried proposers on what they thought was the likely punishing threshold for the anonymous responder. This I did with role-playing and demonstration, as I had not yet fully 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 responders 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, they tended to say that they would not have accepted offers below the one they had made.

The interpretation that proposers make poor estimates of the magnitude of the consequences of their ‘bad’ behavior is reconciled with ethnographic data, which reveal 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 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’). Despite all this, when asked to introspect about the likely social consequences to someone reputed to be a ‘bad person’, my informants consistently drew a blank, no matter how hard I tried to tease it out of them (see Gil-White 2003 for the ethnographic details)

We are faced here with a string of puzzles. First, Torguuds seem to worry a great deal about their reputations as good, honest, upstanding members of society, in great disproportion to the level of social punishment that actually follows from being thought of as bad. Second, in interviews Torguuds appeared quite aware that ‘bad people’ suffer no significant social sanctions—so their fears of reputation loss do not result from a faulty construal of the empirical facts of

⁴ 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.

punishment in everyday life. And yet—and this is the third puzzle—proposers in the ultimatum game *did* seem to misperceive the likelihood of punishment. They reasoned that they would be punished for low-balling the responder (not true), and that they themselves would have punished such offers (probably not true). The puzzles do not appear irresolvable (Gil-White 2003 speculates on one solution).

The other main question of interest in these results is: 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 the minimal group paradigm literature.

The putative importance of the ‘ingroup favoritism’ finding in minimal groups rests depends on this being informative about the real world: behaviors involving significant stakes in the context of real social categories (Tajfel & Turner 1979 argue this case). But if the result obtains only with minimal groups, then it might be a very weak effect that gets swamped by other biases in the context of more realistic situations, in which case we have no more than an interesting exotic result, relevant only to the exotic experimental context of minimal groups, and hardly useful to explaining real discrimination and intergroup conflict.

I shall begin with the issue of stake size. Mummendey et al. (1992) found that participants were scrupulously fair in their allocation of a high-pitched blast of noise. Why? Perhaps because the size of the reward is arguably significant (as opposed to symbolic), and it is also negative, so people may not want to give anonymous others more than their fair share of a sizeable punishment even when they are members of the meaningless outgroup—perhaps because doing so would represent a cost to one’s perception of self as a ‘good person.’ With small or symbolic positive rewards both the ingroup and the outgroup recipient come out with a profit, and since the rewards are absurdly small anyway, it may cost the allocator little or nothing in self-esteem to favor the ingroup recipient slightly. If so, what Mummendey introduced, then, is a manipulation that forces allocators to *pay* (in self-esteem) for their discrimination.

It is possible that sufficiently high and *positive* rewards (by the standards of experimental economics—e.g. \$100) will similarly reduce or erase the size of the allocation bias in minimal group studies. After all, when the stakes are high, the same percentage difference in allocation between ingroup and outgroup recipient constitutes a larger and larger opportunity cost for the outgroup recipient, which then may look more and more to the allocator as unfair punishment that compromises the allocator’s self-image. To my knowledge, however, nobody has tried this.

What should be the effect of increasing the reality of the social categories? From the perspective of theories built around the ‘in-group favoritism’ finding (see Tajfel & Turner 1979; Turner et. al. 1987), one expects the ingroup bias to be stronger with more realistic groups. So in the UG with

ethnicity manipulation there should be a bias to favor the responder when s/he is a coethnic, who will then 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. *And yet the data goes the other way.*

One might argue that since favoring the ingroup responder forces the proposer to take less money in that game, the proposer must *pay* to discriminate. If the effect is strong enough, this may contribute to attenuate and perhaps even erase any tendency to favor the ingroup over the outgroup (as in the Mummendey et al. 1992 study). But this is insufficient to explain the data, given that we have what looks like an *outgroup* favoritism bias. Why does the result go in the opposite direction?

One possible answer is that ethnic groups are norm-groups (see Barth 1969) and in human cognition they are *naturalized* norm-groups, where the differences in normative behavior are naively held by ordinary people to follow from an inner *essence* (innate predispositions) rather than learned behavior (Gil-White, 1999; Gil-White, 2001a; Gil-White 2002). So Torguud proposers may reason that, sharing a common essence with other Torguuds, they have a good guess of how low they 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 may feel less equipped to guess. If proposers are using what they assume to be the ingroup threshold as an anchor, their uncertainty about the outgroup threshold may be sufficient to generate higher offers to the outgroup.

This line of thinking agrees with work done by Zeelenberg & Beattie (1997). Their goal is to explain why offers in the UG are usually considerably higher than observed rejection thresholds. Their data supports their contention that proposers are motivated, at least partly, by a desire to avoid feeling regret—in fact, people find states of affairs that merely contain the *possibility* of regret as aversive. As they say, “in a normal ultimatum game the regret minimizing option is offering too much” because a total opportunity cost (when the responder rejects) is worse than a partial one (when the responder accepts, suggesting that perhaps s/he might have accepted a lower offer). Thus, if proposers (1) feel that they cannot predict where the punishing threshold of the outgroup might be; (2) use what they believe is the ingroup punishing threshold as an anchor; and (3) are motivated to avoid feeling regret, then they will tend to make higher offers to the outgroup.

Not every participant who showed a bias was articulate under cross-examination about why the outgroup responder 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.”

The two main categories of introspections are not necessarily mutually exclusive. Take, for example, the player who made the only *hyper*-fair offer. 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. If ethnic groups are norm groups, and our psychology has evolved to deal with this (Gil-White, 2001a), then insecurity about the punishing threshold is more likely with the outgroup responder. Adding a taste for money and an aversion for regret then predicts higher offers to the outgroup responder.

Another plausible explanation for the result is an anticipation that outgroup members will be likely to punish low offers *coming from an outgroup member*. However, nobody gave such an explanation.

This, of course, does not mean that certain ingroups are not favored discriminatorily, nor does it mean that ethnic groups are not among the groups that more easily generate such ingroup favoritism. Evidence for both propositions is already overwhelming from the historical and current data on nationalism. What the present 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 in ecologically valid social categories.

Indeed, the work of Yamagishi and colleagues suggests that the ‘ingroup favoritism’ bias in minimal groups has nothing to do with discrimination and everything to do with ingroup-cued *reciprocity*. Yamagishi et al. (1998) have found that the ‘ingroup bias’ occurs only when allocators expect in turn to be in the recipient’s role, and thus allocated *to* by anonymous members of the ingroup and outgroup. If they don’t have such an expectation, they allocate to ingroup and outgroup recipients without any bias. So perhaps the observed bias results from an interesting generalized reciprocity norm that is activated only with respect to members of one’s own social category (see also Gaertner & Insko 2000).

Nagging questions remain, however, and the minimal group results continue to defy a straightforward interpretation. If the bias results from ingroup-cued reciprocity, why does the effect disappear when a third group is added (Hartstone & Augoustinos 1995)?

IV. Conclusion

The main methodological lesson for economic experiments in the field is, I think, that we have to be extremely careful about cultural differences that lead to differences of interpretation in our subjects about what the experiment is supposed to be.

An experiment is a *social situation*, replete with social actors who must know their roles and must share a set of background assumptions in order to coordinate with each other. The usual assumption is that subjects in the Western world tend to be university undergraduates who know the routine quite well. Even before their first experience in an experiment, they have heard enough about them that the purpose and context of such social rituals is something that they have well inscribed in their cultural repertoire. When dealing with a non-Western, non-industrialized population, however, we need to be careful to build for our subjects all of those cultural background assumptions about the nature and purpose of experiments, the role of the participant, the role of the experimenter, the role of the institution sponsoring the experiment, etc., etc., etc. Otherwise, our participants creatively imagine a completely different set of background assumptions for the procedures they are helping us carry out.

Perhaps another lesson might be, as suggested by Jonathan Baron (personal communication), that perhaps we should not be so sure that university undergraduates really are so-well socialized that misunderstandings of the sort I found (but perhaps not as extreme) are not common, and merely go under the radar because we are not looking for them.

The present study also shows the value of pilot work. The field experimenter must plan for initial experiments to fail and build in a second one as part of the design. The first experiment should be thought of as a *school* for how to do *that* experiment in *that* field site. It is a lot of work, but a well-conducted field experiment, properly set in its ethnographic context, produces invaluable data that we cannot get in the lab.

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.

References

- Barth, F. (Ed.). (1969). *Ethnic groups and boundaries: The social organization of cultural differences*. Boston: Little Brown & Co.
- Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton, NJ: Princeton University Press.
- Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology*, 1, 263-292.
- Epps, T. W., & Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistics and Computer Simulation*, 26, 177-203.
- Gaertner, L. A., & Insko, C. A. (2000). Intergroup discrimination in the minimal group paradigm: Categorization, reciprocation, or fear? *Journal of Personality & Social Psychology*, 79, 77-94.
- 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(5), 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(4), 515-554.
- Gil-White, F. J. (2001b). A good experiment of choice behavior is a good caricature of a real situation. *Brain and behavioral sciences*, 24(3), 409.
- Gil-White, F. J. (2002). The cognition of ethnicity: Native category systems under the field-experimental microscope. *Field Methods*, 14, 170-198.
- Gil-White, F. J. (forthcoming 2003) Ultimatum game with an ethnicity manipulation: Results from Khovdiin Bulgan Sum, Mongolia. In J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr, and C. Camerer (Eds.). *Foundations of Human Sociality: Ethnography and Experiments in 15 small-scale societies*. Oxford: Oxford University Press.
- Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology*, 25, 179-193.
- Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga. *American Economic Review*, 90(4), 973-979.
- Henrich, J., Boyd, R., Bowles, S., Gintis, H., Fehr, E., and Camerer, C. (2003).

Foundations of Human Sociality: Ethnography and Experiments in 15 small-scale societies. Oxford: Oxford University Press.

Mummendey, A., Simon, B., Dietze, C., Grünert, M., Haeger, G., Kessler, S., Lettgen, S., & Schäferhoff, S. (1992). Categorization is not enough: Intergroup discrimination in negative outcome allocation. *Journal of experimental social psychology, 28*, 155-144.

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.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*(5), 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.

Yamagishi, T., Jin, N., & Miller, A. S. (1998). In-group bias and culture of collectivism. *Asian journal of social psychology, 1*, 315-328.