

# Role-assignment Algorithm: Experimental Study of Cooperation in Voluntary Contributions Public Goods Game and Coordination in Assurance Game

Sun-Ki Chai\*   Kyle Hampton†   Dolgorsuren Dorj‡   Ming Liu§

## Abstract

This study introduces a methodology that builds a cultural profile of subjects using survey questionnaires, then uses the results to predict variations in behavior in controlled computer-mediated experiments across a range of conditions. 156 subjects were provided with pretest surveys, which were used to generate measures for each subject along grid and group, two very prominent and general cultural dimensions drawn from social/cultural anthropology. Grid was hypothesized to induce enforcement of social norms of reciprocity, and group to induce altruism towards other individuals. These subjects were then placed in incentivized computer-mediated repeated interactions involving ten rounds of a social dilemma-inducing Voluntary Contribution Mechanism (VCM) within teams of four. Treatments were varied according to whether team membership were shuffled between stages, whether information was provided on team member's contribution, and by whether clustering of particular types were arranged through certain algorithms. New algorithm differs from regular VCM in such a way that the grid/group characteristics of individuals affect group membership. Overall, it was shown that the group attribute was positively and significantly correlated with the level of individual contribution across all treatments. Individuals were then placed in a modified VCM in which team members were able to punish, at cost to themselves, those who failed to contribute. It was shown the high-gridness individuals punish more than high-groupness individuals, and significantly so for treatments in pooled data. Further studies of grid and group with assurance games are briefly reported and analyzed. It is thus demonstrated that general cultural dimensions can be used to predict differences in behavior under widely varying conditions.

JEL classification codes: C7, C91, C92, D64, M14, Z1.

---

\* Corresponding author Department of Sociology, University of Hawaii at Manoa, 2424 Maile Way, Honolulu, HI 96822. Email: sunki@hawaii.edu

†Department of Economics, Ohio State University, Columbus, Ohio. Email: kyle.hampton@gmail.com

‡Department of Economics, University of Hawaii at Manoa, 2424 Maile Way, Honolulu, HI 96822. Email: dolgorsu@hawaii.edu

§Department of Economics, University of Hawaii at Manoa, 2424 Maile Way, Honolulu, HI 96822. Email: mingliu@hawaii.edu

# 1 Introduction

The role of culture in collective human performance is a natural object of study in this age of globalization, with multinational workforces increasingly becoming the rule in both the private sector and government. Within the military, the necessity for individual soldiers to work with multinational coalitions is becoming more and more prevalent. Moreover, there is a burgeoning literature within the social sciences that studies the effects of culture on collective action. Theories of this sort can be found throughout the social sciences. In economics, it can be found within the rising fields of behavioral economics (Camerer, Loewenstein, and Rabin 2004; Thaler 2007; Diamond 2007; Wilkinson 2007) and experimental economics (Friedman and Sunder 1994; Friedman 2004; Kagel and Roth 1997), where increasing notice is made of individuals who do not behave as uniform, homogenous rational maximizers. In other areas of economics, examination is made of the role of morality in market behavior (Zak 2008; Shermer 2008), social exchange and status (Gintis et al 2006), choice over time (Loewenstein 2003), and of social learning (Chamley 2004). Behavioral game theory is taking increasing notice of reputations (Mailath and Samuelson 2006) and learning (Cesa-Bianchi 2006). Finally, notice is being made increasingly of the cultural variations across countries and societies in the way that individuals behave under the same sets of circumstances (Henrich et al 2004). Computational economics (Holland 1998) and artificial life studies (Epstein and Axtell 1996; Epstein 2007) have always claimed to incorporate cultural factors into their analysis, but their rule-based models do not always make clear from where cultural determinants of rules arise.

One indication of the increased role of culture in social and behavioral theorizing is the tendency for various fields of study to blend together, such as recent collaborations between economists and psychologists (Frey and Stutzer 2007; Cremer 2006; Brocas and Carillo 2003). In sociology, the field of cultural sociology (Smith 1998; 2001; Alexander 2003) is undergoing a renaissance by being brought together with psychological theories into the new field of cognitive sociology (Zerubavel 1997; Cerulo 2002), as well as the more long-lasting field of sociological social psychology (Burke 2006), which is influenced by economic choice models as well as psychological ones. Finally, the study of acculturation psychology brings sociological analysis of norm formation into the field of cross-cultural psychology (Sam and Berry 2006).

Given all this intellectual activity revolving around questions of culture and collective action, one would think that there would be within this activity considerable attention paid to the role of culture in collective performance. Unfortunately, this is not the case. In a recent survey of theories of collective performance (Stock 2004), it is evident that cultural factors are examined only very marginally in analysis of performance, and the interaction of cultures virtually ignored. There may be a number of reasons for this, but perhaps the greatest may be the lack of an appropriate framework on which to base such an analysis.

The paper addresses this shortcoming by incorporating the role of two cultural factors in particular, grid and group, the basis for perhaps the most prominent general framework for cultural classification in the social sciences, and does so in a way that shows how grid and group interact with the organizationally-induced structure of collective action to determine individual and collective performance.

This study empirically demonstrates how sorting of particular types of individuals based on grid and group cultural attributes may increase performance in, systematic, predictable ways, in a range of settings, making use of the methodology of computer-mediated experiments. This is the first study of any kind to ever show that cultural characteristics, measured in quantitative form, can account for predictable differences in behaviors under controlled experimental conditions. In particular, it shows how clustering of high-groupness subjects improves overall organization output. It shows how under certain institutional structures organization may benefit from high-gridness individuals' actions. We simply aim to provide a better understanding and use of cultural factors to overcome social dilemmas in organizations that have team structure.

More specifically, first, the survey questionnaire were used to generate measures for each subject along the grid and group. Second, we study cooperation in teams through algorithms that sort particular cultural types in voluntary contribution mechanism (VCM) with and without punishment: (i) in the algorithm that uses VCM without punishment the group attribute affects the group membership; (ii) information on others contributions was manipulated in such a way that only contributions of high-groupness individuals were seen by others in VCM with no punishment; (iii) we cluster high-gridness individuals with low-groupness individuals in the VCM with punishment; (iv) we assign only high-gridness types in the punisher's roles in the VCM with punishment and compare with situation where only low-gridness individuals have a punisher tasks. In addition, we examine how clustering of high-gridness individuals affect coordination in the assurance game.

Our results show that cultural dimensions such as grid and group can be used as a predictors of behavior under the team organizational structure. In particular, clustering of high-groupness individuals in one team may increase overall performance of organization that consists of multiple teams. Note that this sorting mechanism generates more homogenous teams. On the other hand, full transparency of the output produced by high-groupness individuals may turn this type into a role model within more heterogeneous team. However, this result was insignificant or produced negligible or no effect. Our sorting mechanism that clusters high-gridness individuals with low-groupness individuals which prevents slack of job by latter type produced inconclusive result. The efficiency in the sorting mechanism that assigns high-gridness individuals into norm enforcer's roles was no different from the efficiency in the random assignment. Overall, our results demonstrate that group attribute positively and significantly affects contributions in VCM without punishment while grid score significantly affects punishing behavior in VCM with punishment opportunity.

The paper proceeds as follows. In the next section we overview the literature on behavioral types sorting. Section 2 provides the formal representation of grid/group cultural theory and explains algorithms that assign cultural types in certain roles. Section 3 provides experimental design. Section 4 presents the results and we conclude in section 5.

## **1. Heterogeneity and Sorting Algorithms**

The repeated VCM has revealed several consistent results regarding the behavior of individuals asked to engage in collective action. In early rounds, subjects tend to make

substantial positive contributions to the public good despite game-theoretical predictions of zero. In later rounds, subjects substantially decrease their contributions, a result that many economists have chalked to subject learning in the course of the game (Andreoni 1995).

More recently however, this idea has come into question as new experiment designs have allowed researchers to devise institutional rules that maintain high levels of contributions throughout the course of the experiment. One of these designs allows subjects to selectively impose financial penalties on other subjects within the group in response to his or her level of contribution to the public good (Fehr and Gächter, 2000). More recently, others have discovered that grouping high contributors with other high contributors in succeeding rounds has the similar effect of sustaining contribution levels (Gunnthorsdottir et al. 2007). However, it is not clear in this literature how to identify ahead of time those who will be cooperators within a particular VCM.

Each of the approaches seems to suggest a complex story of what is motivating both the high levels of initial contributions to the public good and the deterioration of those contribution levels in more traditional VCM formats. In particular, it has been observed that there may exist substantial heterogeneity in the attitudes of human beings regarding collective action and that the decrease in contributions to the public good may have less to do with subjects learning the nature of the game and more to do with subjects learning the nature of the other subjects with which they are interacting.

This distinction has led many researchers to posit the existence of distinct types of individuals. For instance, it has become de rigueur for researchers to identify three distinct types of players in the VCM: unconditional cooperators, unconditional noncooperators, and reciprocators who typically begin by contributing to the public good in early rounds but condition their later contributions on the contributions of others within their group (Kurzban and Houser, 2005, Fishbacher et al. 2001).

This grouping of individuals into distinct types has allowed for an entirely different approach to examining human behavior in VCMs. The researchers identifying these types suggest that the observed decay in contributions is a result of reciprocators reacting to the decisions of unconditional non-cooperators. This suggests the potential for studying what Ones and Putterman (2007) refer to as an ecology of interacting types. The use of the term ecology is apt as distinctive outcomes seem likely to arise from the complicated interplay of heterogeneous personalities and the institutions through which these interactions take place.

Research into these dynamics to this point has focused primarily on the phenomenon of sorting, where it is presumed that cooperators, both conditional and unconditional, are likely to interact only with one another while excluding unconditional non-cooperators. It is an intuitive scenario and one observed in various environments where collective action takes place. Experimental economists have simulated this environment by implementing a traditional VCM where subjects are matched with new group members each round. However, they alter the traditional format by grouping subjects on the basis of past contributions to the public good. High contributors are matched with other high contributors. Low contributors are matched with other low contributors.

The result of this sorting is an increase in the overall level of contributions as conditional cooperators are no longer compelled to reduce their contributions

in response to the actions of unconditional non-cooperators. Similar dynamics have been observed when sorting subjects in both the prisoner's dilemma experiment (Bohnet and Kuebler, 2005) and in an extensive form "trust" game (Rigdon et al., 2007).

At the heart of this sorting process is an algorithm devised by the experimenter to translate a single input, the past contributions of the subjects into the group account, into a group assignment. As such these algorithms have been simple and straight forward, meant primarily to demonstrate the existence of a sorting rule that will result in an increase in social surplus and overall performance of team.

Our primary research goal is to extend this line of inquiry by devising and testing a more comprehensive algorithm that not only includes an increase in input variables but a distinct change in the nature of those inputs. As opposed to an endogenous determination of these inputs by recording the past subject decisions in the experiment, we plan to utilize a survey instrument to elicit the subject type prior to any decision on their part.

The survey instrument used, the grid-group survey, has several distinct advantages over the heretofore standard of using past behavior to sort individuals into groups. The first is primarily practical. Oftentimes, the nature of a collective action problem precludes the use of past behavior as a determinant of grouping. The motivating example for this research is in fact teamwork in a corporate setting where a secondary reallocation of individuals to different groupings simply may not be feasible. If effective, a survey can provide a relatively low cost means of sorting individuals prior to actual collective action.

But more importantly, the grid-group survey instrument allows one a more nuanced definition of the type of individuals being allocated. As mentioned above, economists have found it useful to classify individuals as one of three types. Two of these types have seem to have clear correlates in the gridgroup framework: unconditional cooperators (High Group - Low Grid) and unconditional non-cooperators (Low Group - Low Grid). The third category, conditional cooperators, would seem to encompass High Grid individuals of either type, high or low group. By distinguishing the grid and group characteristics of the individuals to be sorted, we can get a more comprehensive understanding of the interaction dynamics of these unique types. Moreover, because the grid-group framework is a general one that applies to multiple kinds of games, not just VCMs, it can help to infer how individuals who behave in a particular way in a VCM might behave in a completely different type of interactional environment.

Moreover, we plan to expand the output of the algorithm beyond merely determining the grouping of various subject types. Our algorithm will also assign "roles" to subjects within those groups, with these roles defined by a differential access to information regarding the public good contributions of others and the ability of other to view your own level of contribution.

In cross-cultural study by Yamagishi and Yamagishi (1988) there was no attempt to directly measure the cultural characteristics of US and Japanese citizens, they make predictions based upon conventional stereotypes about the differences between the two cultures. According one view, Japanese should prefer collective values and trust more while American should be individualists and therefore trust less. However, lab experiments provide opposite evidence. In a no-punishment condition, Japanese

individuals contribute less (44 percent of their endowment) and Americans individuals more (56.2 percent of their endowment) to a collective enterprise. However, when sanctions were available, cooperation was high across both subject pools (74.6 and 75.5 percent respectively). This suggests that Japanese individuals cooperate often within the group not because of an intrinsic desire to do so, but instead because of the presence of social norms that encourage sanctions against those who fail to do so. Note, however, that imposition of such sanctions itself is costly, and therefore cannot be explained by merely self-interested behavior. Grid-group theory can not only help explain these differences between Japanese and Americans, but can show how they can be generalized to other countries and other types of interactions. Moreover, it shows how it is possible to go beyond stereotypes of national character and examine cultural differences at an individual and sub-national level. Besides showing how grid and group affect individual performance in different types of organizational settings, we also show that with heterogeneous population it is possible to improve aggregate performance through the grouping of certain types of individuals in particular roles. Specifically, we test importance of cultural factors in sustaining high performance; for example, high-grid (strong reciprocity) and high-group (altruism) characteristics are among the explanations of shift in contributions over repeated interactions.

## 2 Grid-Group Formalization and Role-Assignment Algorithm

This study is designed to test a model that uses grid and group to map individuals to particular roles within a team, examining the effect of that mapping on aggregate team performance. This section will provide an overview of the individual-level grid-group formalization, as well as the role-assignment algorithm. A team  $T$  comprises an exogenously determined number of individuals. There is a corresponding set of task roles  $I$  such that  $\|I\| \leq \|T\|$  that must be filled by some subset of the team members. Let us define  $\phi: I \rightarrow T$  as a function mapping roles onto specific individuals or subgroups within the team. It is the culturally-based role assignment algorithm seeks to determine  $\phi$  in a way that provides improvement in performance yet maintains generality, calculational tractability, and compatibility with available data. Each individual  $s \in T$  has the characteristics of grid ( $m_s$ ) and group ( $n_s$ ), s.t.

$$\forall s \in T : m_s, n_s \in [0,1]$$

Broadly speaking, grid has been defined as an attachment to rule-following determined by institutionalized social norms and group has been defined as incorporation of identity into a larger collective unit (Douglas 1970; 1982). In the process of being formalized for the purposes of incorporation into quantitative predictive models (Chai and Wildavsky 1994; Chai and Swedlow 1998), the grid has been further defined more concretely and in choice-theoretic terms as expressive utility placed on following social norms defined within a collectivity, and group as altruistic utility towards members of the collectivity. This leads us to the following transformation of individual utility functions based on gridgroup cultural traits:

With  $x_s$  is defined as individual  $s$ 's (one period) untransformed utility. A individual's group-transformed utility  $y_s$  is specified as linear altruism

$$y_s = \left( \sum_{t \supset s} n_s x_t \right) + x_s \quad (2)$$

Let us define  $a_s^*$  as the socially prescribed action for the role into which individual  $s$  has been placed. An individual's grid and group-transformed utility is then represented as:

$$z_s = y_s (\text{ord}(a_s = a_s^*) + (1 - m_s) \text{ord}(a_s \neq a_s^*)) \quad (3)$$

It should be noted that a simplifications have made here from a full translation of grid/group theory, primarily in light of practical data limitations, and to a lesser extent tractability. Among other things, the groupness and gridness variables  $n_s$  and  $m_s$  may apply differentially to members of a collective, depending on the extent to which certain individuals are judged to be part of an individual's in-group". However, generating this information with a degree of subtlety and validity in experimental setting is problematic. A related issue is that groupness and gridness can both conceivably be negative, i.e. an individual may have negative altruism (misanthropy) towards others or may gain expressive utility from violating norms. However, the originally grid-group theory posits that such values cannot proliferate in any society that is capable of surviving. Nonetheless, it is preferable to reintroduce such variables at some point, and a short discussion of how this may be done follows at the end of this document. Nonetheless, it is preferable to reintroduce such variables at some point, and a short discussion of how this may be done follows at the end of this document.

Grid and group data can be derived from a variety of sources. Ideally, data collection will take place directly, from the specific individuals being modeled. If data collection is done at this individual level, then the most likely sources of information are surveys and experiments. Surveys will typically consist of questions about basic values and ideas held by individuals and/or about the decisions they would make under a range of controlled (usually generic) scenarios. Experiments will place subjects in simulated versions of such scenarios, with real rewards at stake, to measure how variations in the subject and scenario affect decisions. In recent years, with the level of control provided by computer-mediated experimental laboratories, ability to use the latter methodology in particular has greatly improved, and can often be combined with electronic pretest and/or post-test surveys to test and verify the effect of grid and group on behavior.

Where individual-level data for members of a team is not available and not feasible to collect due to time or other constraints, it is possible to obtain grid-group values (though with some obvious loss of accuracy) by associating them associated with a cultural collective (nation, ethnic group, community)  $T_\epsilon \in T$ . Such collectives may be treated as having uniform culture, i.e. the same group grid and group, across all their members, which in turn will be assigned to each individual in a collectivity. Even where some individual-level data exists, it may still have to be aggregated if this data does not extend to all members of members of the team. The size of the collectives being considered will depend in part on the level of aggregation at which data is available, but also upon the boundaries that match most closely with the formal criteria that are used to divide the team into units. As mentioned, a role-assignment algorithm attempts to best match individuals in  $T$  with roles in  $I$  based on matching their attributes. Each role  $i \in I$  is characterized by the following attributes:

$\mu_{ii} \in [-1;1]; i \in I$  : the extent to which i's role requires exercise of discretion and improvisation, as opposed to the implementation of clearly specified and unambiguous rules and procedures,  
 $\mu_{ij} \in [-1;1]; i \neq j \in I$  : the extent to which i's role requires real time coordination with j,  
 $\nu_{ii}; i \in I$  : the expected direct net benefit that is gained by an individual i's performance of their task role, and  
 $\nu_{ij}; i \neq j \in I$  : the expected direct net benefit gained by j from i's performance of task role.

The exact scale of the  $\mu$  and  $\nu$  values is not crucial as long as it is consistent across individuals in the way that that data are derived for analyzing a particular team and task. In many cases, especially for large teams, it is likely that multiple task roles may have identical characteristics, i.e. that there are types of tasks for which a single type must be filled by multiple individuals. In that case, the user may specify each type of task and the number of individuals that are necessary to occupy it, rather than each individual task separately. Similarly, one might have identical  $\mu_{ij}$  for multiple values of  $j$ , and likewise for  $\nu_{ij}$ .

There may be a variety of ways in which the  $\mu$  and  $\nu$  data may be collected. Particularly for tasks that are not formally codified, subjective coding may be used. Information for the each value of these variables can be provided by subject matter experts for the task domain of interest. This can be done through a straightforward survey that ask the subject matter experts (SMEs) to enter numbers into uniform subjective scales, or through a more sophisticated elicitation methodology, perhaps combined with an experimental setting. However the opinions are elicited, they in turn will be used as a "databank" once the user decides which of these specific tasks within a domain are pertinent to the collective assignment at hand.

Alternatively, where we are dealing with abstract tasks or where tasks are highly codified,  $\mu$  and  $\nu$  may be derived as an automated mapping from the formal description of the tasks. The way in which this will be done would naturally depend on the nature of the formal description. In the case where the nature of tasks can be represented using the formalizations of game theory (as in the experiments reported elsewhere in this report) a mapping can be drawn by examining at the payoff and information structure of this game.

The inclusion of these independent variables reflects both personal and interpersonal effects of grid and group. Grid's personal effects reflect its definition as the extent to which culture tends to bind its members to defined procedures and routines, which cause individuals in high grid cultures to generally be better suited for tasks that are predictable, as opposed to those requiring improvisation, while the converse is true of individuals in low grid cultures. Its interpersonal effects reflect the fact that clearly defined procedures tend to make it easier for high grid individuals to act collectively.

Group's personal and interpersonal effects are more uniform in nature, and reflect the relative weights that an individual places on their own payoffs and the payoffs of others in their group. In the economic literature, this is often referred to as "altruistic" preferences, and this can be reflected in terms of a utility function" that reflects the relative desirability of different outcomes.

For each specification, the system generates a performance function  $u_i$  for each role  $i$ . There is no real precedent in the academic literature for specifying the functional form for this performance function, since the whole idea of calculating numerically the impact of culture on human performance is a very novel one. The following is presented as a reasonable approach given what is known about grid and group:

$$u_i = (\mu_{ii} + m_{\phi(i)} \sum_{j \neq i} \mu_{ij})(v_{ii} + n_{\phi(i)} \sum_{j \neq i} v_{ij}) \quad (4)$$

The overall performance metric can be specified simply as  $Y = \sum_{i \in I} u_i$ . The job of algorithm, then, is to determine a mapping  $\phi$  that provides the highest value of  $Y$ . This is the role assignment that is then recommended to the user. This formula captures the fact that individuals with high gridness are suited for tasks that require high degrees of coordination with other individuals, since it is more likely that they will be able to converge their behaviors around existing social norms. Likewise, it captures the fact that individuals with high groupness are suited for tasks in which their actions have sizable consequences for the well-being of other individuals, since they will be more likely to incorporate these consequences intrinsically into their decisions.

We of course cannot claim a priori that there is no performance formula better than this one, only that this is grounded and derived in a systematic fashion from established theory, and thus can be expected to perform well compared to random or less systematic assignment rules. As mentioned, if data is available, there are additional variables that could be incorporated into the formula in order to potentially improve its predictive accuracy. Even with the existing variables, the usefulness of this formula, like all general social science models, will vary from context to context, and multiple versions may be necessary for better performance across such contexts. Nonetheless, this formula is seen as a reasonable starting point, one that can be empirically demonstrated to have value and can be a starting point for the evolution and development of improved performance formulas and models.

## System Use and Technical Constraints

The main use of the envisioned decision support system is for the user to specify the make up of the team, as well as the duties of each role, and for the system to optimally assign team members to each role. This involves an generating an ideal assignment function  $\phi^*$  that can maximize  $Y$ . One issue is that the optimization problem does not admit a simple solution concept, and an exhaustive search algorithm would be appropriate only for teams with relatively small numbers of roles. It may thus be necessary to simplify the performance metric in some way to make the determination of optimal assignment more tractable, while keeping loss of accuracy down.

In practice, the grid-group scores for members of the team will have preferably been from primary data collection at the individual level, either from surveys, experiments, or a mixed-methodology. Where this is not feasible, they can be inferred by taking means from secondary data collected earlier from other individuals of their same national or other communal affiliation (depending on availability). The disadvantage of the latter, of course, is that it does not allow us to look at the effects of variations within the group.

One additional feature of the system, other than those aspects that deal directly with culture and roles, is based on the fact that there will be certain organizational limitations on assignments of individuals to roles. There will essentially be two kinds of such constraints. The first is that certain kinds of roles will require particular technical skills not possessed by all members of the team. The second constraint is that there may be already existing units within the team, which for various reasons cannot be split up into different roles. This would be more likely for larger teams that have been recently assembled from different organizations, with examples being multinational peacekeeping forces, investment partnerships between two firms, and temporary consulting arrangements.

While in some cases there may in reality be skills that are useful but not necessary” or existing unit boundaries that can but breached with some disruption, incorporating such facts into the system would introduce additional complexity, and would require a difficult subjective comparison between cultural and non-cultural costs that would be difficult to make. Hence, specifying technical skills in terms of constraints is probably necessary for a relatively manageable system.

Such limitations will need to be specified, and the system will simply not consider combinations that are outside of such limitations. This will actually tend to simplify calculations, since the combination of possibilities that the system needs to consider will be diminished.

Each role and unit’s characteristics and organizational limitations are specified by the user through a query front end. Naturally, part of the emphasis in developing this front end should be on making the process as painless as possible, which in turn may mean asking questions in an indirect and more qualitative way than specified here, then translating the answers into the requisite numbers. This may even improve the validity of the resulting numbers.

## **Operationalization of the Models**

### **Operationalization of Grid and Group**

We conducted a survey before each experiment to obtain cultural characteristics. Then we quantify the cultural characteristics of each individual based on answers they pursue during a survey that consists of 29 questions (see appendix), which transmits information on personal values with respect to following key elements: (i) social values, (ii) opinion on employment, (iii) national goals, (iv) social distance, (v) workplace, (vi) religion and (vii) social norms. See survey attached for details. All subjects were ranked according to grid/group characteristics. First, we calculate grid<sup>1</sup> and group<sup>2</sup> indexes in line with the formula in the footnote. This method for calculating grid and group is determined by

---

<sup>1</sup> Grid= $((4-\text{Answer}[10])/3+(3-\text{Answer}[13])/2+(3-\text{Answer}[14])/2+(3-\text{Answer}[17])/2+(2-\text{Answer}[18])/1+(3-\text{Answer}[22])/2+(\text{Answer}[25]-1)/9+(10-\text{Answer}[26])/9+(10-\text{Answer}[27])/9+(10-\text{Answer}[28])/9+(10-\text{Answer}[29])/9)/11$ ;

<sup>2</sup> Group= $((4-\text{Answer}[8])/3+(4-\text{Answer}[9])/3+(2-\text{Answer}[11])/1+(2-\text{Answer}[12])/1+(\text{Answer}[15]-1)/1+(\text{Answer}[16]-1)/3+(3-\text{Answer}[19])/2+(3-\text{Answer}[20])/2+(\text{Answer}[21]-1)/2+(\text{Answer}[23]-1)/9+(10-\text{Answer}[24])/9)/11$ ; The Answer[1], Answer[2] are the numbers that quantify and identify the exact answers provided by subjects in the survey.

taking selected items from the World Values Survey (Inglehart et al. 1998, 2004). It has an advantage of allowing the survey results to be verified on a national basis against the much larger sample that the WVS provides. For reasons of robustness, two additional questions were placed on the survey, each of them being original items that seek to encapsulate in a single indicator the qualities that make up grid (Answer[7]) and group (Answer[6]).

### **Operationalization of Untransformed Utility and Social Norms**

Even with the full specification of the individual-level grid-group and role assignment models, there are matters of interpretation in how the variables of models will be operationalized in a concrete situation or test. For grid and group, it is necessary to specify the nature of untransformed utility and relevant social norms within the situation being modeled. There is really no way to do this in a fashion that is universally applicable given a finite amount of research resources. There is a potentially infinite number of goods and outcomes that individuals in different societies might potentially care about as a matter of utility. Likewise, the number of social norms that exist in any single society, much less across all societies, is too large to be countable.

Instead, the nature of untransformed utility and social norms must either be based on measurements that are specific to a situation or frame (when applied in concrete, realtime settings) or to goods that are valued positively and rules that are known across a wide range of societies and generally applicable to the full range of choice environments (when applied in controlled experimental settings). In the former case, one would require the sustained input of a set of subject matter experts who are familiar both with the nature of the task and the nature of each of the cultures that are participating in the task. The enrollment of such SMEs for a non-trivial number of cultures is beyond clearly beyond the scope of this paper. Indeed, one of the advantages of computer-mediated experimental conditions, besides the degree of control that it provides over information and incentives, is that the laboratory conditions help to filter out the innumerable difficult-to-measure idiosyncratic factors that characterize natural conditions but also make it difficult to attribute the causes of difference in behavior (Friedman 2007).

In the lab, it is possible to reduce the possible goods that individuals might plausibly be concerned with in regards to their culture, regardless of culture, to the monetary payoffs that they and their fellow team members receive from the experiments. Other factors, such as social prestige, love, aesthetic pleasures, etc. are minimized by the anonymity of the interaction and the abstract nature of the tasks. Likewise, it is possible to reduce the types of norms that apply to those that relate to fair or just treatment of other individuals with regards to such payoffs based on those individuals past behavior, i.e. to norms of reciprocity (Andreoni 1995; Falk and Fischbacher 1998; Bolton and Ockenfels 2000).

Note that limiting relevant goods and norms in this fashion, while obviously simplifying reality, retains a decisive role for culture in determining individual behavior. Group will decide the extent to which an individual will value the payoffs of their fellow team members vis-a-vis his/her own, while grid will determine an individual will feel bound by norms of reciprocity, even at a cost to their own personal payoff.

## Operationalization of coordination and dependency

In the experiments, the structure of information and incentives associated with each role will be encapsulated by a standard game theoretical formalization. This will promote abstract representation of a general range of problems, as well as technical clarity regarding the decision calculus facing each subject. This makes it possible to isolate the effect of each variable. As the preceding literature reviews show, this is a reason why such an approach has been dominant in social science computer-mediated experiments. As mentioned, another advantage of the game theory formalization is that it provides with a basis for determining the characteristics of roles without resorting to subject SME judgement. Each player in a game is assigned a formal role, one that carries with it a set of choices, each with its own information set, as well as a payoff function that determines the outcome for each player as an outcome of all players' moves.

That being said, the translation from game-theoretic notation to concepts that are amenable to cultural characterization is not straightforward and requires a certain amount of theoretical inventiveness. As with the role assignment formula itself, there is no way to determine a priori, what is the best operationalization from a game structure to the major role constructs of coordination and dependence. Furthermore, there is clearly no universal operationalization that will be tractable and lead to optimal role allocation in all cases. Instead, the operationalization should provide a justifiable heuristic that translates game structure into culturally relevant constructs. And as with any first-cut attempt at modeling a novel domain, the initial operationalization should be a simple one.

Perhaps the most simple and intuitive way of determining  $\mu$  and  $\nu$  for a game structure is, in a rough formalization, focusing on mean value and variation:

$$\forall i, j : \nu_{ij} = E_A(x_j(b, a) - x_j(c, a)) \quad (5)$$

$$\forall i, j : \mu_{ij} = Var_A(x_j(b, a) - x_j(c, a)) \quad (6)$$

For simple dichotomous choice environments, where  $b$  and  $c \in A_i$  are the two choices of actions such that  $\sum_j E_A(x_j(b, a) - x_j(c, a)) \geq 0$  and  $a$  is a derived probability distribution over feasible actions of all players other than  $i$ . For more complex choice environments,  $b$  can represent  $\text{ArgMax}_{a \in A}$ , and  $c$  can represent  $\text{ArgMin}_{a \in A}$ . By taking expected values of payoffs, we capture the magnitude of the effect of an individual's actions on her own payoff and that of others, consistent with the conceptualization of interdependence above. By taking the variance, we capture the extent to which this effect depends on the actions of others, consistent with the conceptualization of coordination need. Hence, while the operationalization is no doubt far from perfect, it has the kind of face validity and symmetry, along with the desired simplicity. By no means is this meant to be a general methodology for operationalizing the role assignment algorithm to formal game structures, as a number of matters of interpretation that remain open. There are number of tricky parts to this, not the least of which is how to assign subjective probabilities to profiles of actions to determine the expected value (E) and variance (Var). In practice, the way that this is done depends on the context in which the modeling is taking place. In the absence of any information about the other players and their cultural types, it would make sense to assign each profile of actions equal probability. But in many of our analysis we may in fact know something about the cultures of other players, as well as the payoff and information pertinent to the role that they have been placed. In cases where individuals

have at least partial knowledge about the nature and role of the other individuals participating in their joint activity, it becomes a kind of equilibrium search problem. For instance, zero probability should be assigned to any profile containing an action that the individual knows is dominated for another individual in the group. Further exploration of the Stage II treatments for the experiments can demonstrate how this works in practice.

### 3 Experimental Design

In this paper we test the role-assignment algorithm under a more complex set of conditions, in which roles become highly differentiated and specialized. In particular, the algorithm provides guides for assignment and organization of a team that incorporates heterogeneous individuals with varying levels of grid and group. We examine the differences in performance of the algorithm-assigned group compared to a randomly assigned control group. We employ several games to capture group performance: VCM with and without penalty point assignment, and assurance game.

Our experiment includes 12 treatments in four designs. Design 1 consists of two treatments in 20 periods: the first ten-period game is an iterated voluntary contribution mechanism (VCM) with no punishment (NP) and sorting; followed with another ten-period game of iterated VCM with punishment (P) and sorting.

Design 2 contains two treatments in 20 periods as comparison of Design 1: the first ten-period in Design 2 is also iterated VCM with no punishment (NP) but no sorting; the following 10 periods is iterated VCM with punishment (P) and also no sorting.

Furthermore, four treatments of Design 3 have 24 periods: the first ten-period games are NP-condition-differentiated group roles with sorting followed by another ten-period game of P-condition-differentiated grid role with sorting. In this design, we also include a two-period 4-person Assurance game and a two-period 2-person Assurance game with controlled assignment.

Lastly, Design 4 comprises of four treatments in 24 periods including a ten-period game of NP-condition-differentiated group role with no sorting and ten-period game of P-condition-differentiated grid role with no sorting; two-period 4-person Assurance game and 2-person Assurance game where random assignment treatments serves as a control for the pre-specified algorithm treatments with sorting. The summary of the design is listed in Table 1 and detailed descriptions for each design and treatment are provided in the following sections.

To ensure the understanding of payoffs involved in the experiment subjects completed a quiz and two practice periods are conducted before treatment starts to ensure that they are familiar with the software. After the experiment, a questionnaire is given to ascertain the participants' strategies and opinions about the experience.

For each of these treatments, we take the goal for the performance function to be the maximization of the aggregate payoff of all the actors in each team. This is not to say that the performance function will perfectly predict aggregate payoff, but that there should be a strong positive correlation between the two.

#### **Design 1**

### Collective compensation with sorting (NPA)(1-1)<sup>3</sup>

*Game Setup:* Partner VCM with no punishment (NP), 10 rounds, teams of four  
*Assignment Rule:* high group individuals are concentrated with other high group individuals as much as possible, leaving the low group individuals in other teams.

An NP-condition reproduces a linear public good game where the number of people in the group is set to four. This treatment maximizes team performance based on the idea that high-group individuals tend to cooperate; at the same time there are many conditional cooperators who will induce high effort if others will exert high effort. Conditional cooperators tend to deviate from cooperation when others free ride and exert low effort. Hence to ensure that conditional reciprocators will not retaliate seeing other free riding we separate low group individuals from high group individuals. This way we maximize the overall welfare of the society/community that consists of several units/teams.

Given the formal definition of group, it is clear that high-group individuals can appropriate a larger portion of the contribution to a public good through their intrinsic valuation of the payoffs of other members of their team. However, what implications does this have for assignment of such individuals? In this type of game, certain individuals can have opportunities to contribute a larger amount to the public good than others.

It is noted that the effects of actions in a classic (NP) VCM are non-coordinative. In other words, the relative effect of an individual's actions on their own and others' utility does not depend on what the others are doing. Hence  $v_{ij} = 1$  for all  $i$  and  $j$ , and has no effect on any single round of a VCM game. On the other hand,  $\mu_{ij}$  will be directly proportional to the size of the contribution that the individual's role specifies for all  $i$  and  $j$ . Thus  $Y$  reduces to  $\sum_{i \in I} (\mu_{ii} + \sum_{j \neq i} \mu_{ij} m_{\phi(i)})$ . In our fairly simple experimental designs,  $\mu_{ij}$  will not differ for any  $i$ , given any particular  $j$ . It is then clear that, all else being equal,  $Y$  will be maximized in general by assigning individuals with high  $m$  (groupness) to roles in which  $\mu$  is highest.

In a partner treatment, the dependence that a particular role contains depends in part on the nature of the other members in the team, i.e.  $\mu_{ij} = E_A(x_j(b, a) - x_j(c, a))$  is higher when action  $b$  (action of public investment) engenders contributing behaviors in subsequent periods from other members of the team. This in turn, according to the implications of grid-group theory depends on the groupness of the other members of the team. Hence clustering of high-groupness members in the same teams can be expected to lead not only to higher team performance, but higher aggregate performance by all the teams in a session.

In this treatment subjects assignment schedule follows as: (i) given the grid/group score for each subject, calculate the average of the scores present in the session with twenty subjects ( $N=20$ ); (ii) now rank all subjects in terms of group score; (iii) assign

---

<sup>3</sup> We denote the design and treatment as (1-1) (1-2), etc. where 1-1 represent Design 1 treatment 1 and 1-2 represent Design 1 and treatment 2 and so on.

first highest ranked high-group four subjects (ranked 1-4) in a first team, then assign next four individuals (ranked 5-8) with highest group score into second team; assign next four highest group scored bundle of subjects (ranked 9-12) into third team, etc. An assignment schedule generates relatively homogeneous teams to maximize overall system performance. This algorithm is similar to Gunthorsdottir et al. (2007) sorting of cooperative subjects in one team and less cooperative ones in other groups. However, their clustering technique based on subject's contributions, while in our study algorithm based on groupness characteristic which was calculated using answers subjects provided in the survey right before the treatments started.

### **Enforced compensation regime with sorting (PA) (1-2)**

*Game Setup:* partner VCM with punishment, 10 rounds, teams of four.

*Assignment Rule:* high grid individuals are distributed in higher numbers in those groups which have low numbers of high-group individuals

A VCM with punishment (P-VCM), is similar to an NP-VCM, except for the fact that, after contribution, there is another move in which individuals may punish others whom they do not think contributed sufficiently to the public good. This punishment is at cost to both the punisher and punishee, so there is no benefit to doing so except to promote a "just" outcome and, in a repeated game, to deter further failure to contribute. Punishment of failure to contribute is a form of strong reciprocity, a social norm that is accepted in by most experimental social scientists to be one (and perhaps the only one) that is universally shared in by all societies. Because this is a more complex game setup than the NP-VCM, there is interaction between the moves of individuals. In particular, an individual can commit to punishing failures to contribute, and the other individuals in the team are aware of this, they will choose to contribute because the cost of receiving punishment exceeds any individual benefits from failing to contribute. Hence the greater the intensity of punishment an individual is able to impart, the greater  $Var_A(x_j(b, \tilde{a}) - x_j(c, a))$ . Hence, though a similar logic to that which recommends assigning high-group individuals to roles having larger contributonal abilities in NP-VCMs, this would recommend assigning high-grid individuals to roles having larger punishment abilities in P-VCMs.

The P-condition differs from NP-condition by the existence of a punishment stage. Here, to ensure that a punishment mechanism (social norms) is followed by everyone, we assign enough number of high-grid individuals (strong reciprocators) to each team. Strong reciprocators are willing to enforce the rules of the team and hence they are helpful to sustain cooperation and high performance<sup>4</sup>.

P-condition has different assignment rule as follows: (i) all subjects are ranked according their group score and grid score; (ii) we rank all subjects according the group

---

<sup>4</sup> The essential feature of strong reciprocity is a willingness to sacrifice resources for rewarding fair and punishing unfair behavior even if this is costly and provides neither present nor future material rewards for the reciprocator (Fehr, Fischbacher and Gächter, 2002).

score and take the half of subjects with lowest group scores within the session<sup>5</sup> as a target. Then we rank the rest of population according the grid score; (iv) in the next step of algorithm two highest-grid scored subjects are matched with two lowest-group scored subjects in teams of four. This way we ensure that all low group scored subjects would not cluster in one team and high-gridness individuals have opportunity to discipline low-group types. Here assignment allows obtaining high performance in a team with mixed cultural types

### **Design 2:**

#### **Collective compensation without sorting (NPR) (2-1)**

*Game Setup:* partner VCM with no punishment, 10 periods, teams of four.

*Assignment Rule:* random assignment.

#### **Enforced compensation regime without sorting (PR) (2-2)**

*Game Setup:* partner VCM with punishment, 10 periods, teams of four.

*Assignment Rule:* random assignment.

### **Design 3:**

#### **Collective compensation-partial visibility-with sorting (NPPA) (3-1)**

*Game Setup:* Shuffle VCM with no punishment, 10 periods, teams of four. One individual's contribution is visible to everyone in periods 1-5, rest not visible, identity individuals who are visible is not announced. Two individual's contributions are visible to everyone in periods 6-10. Total payoff cannot be seen until end.

*Assignment Rule:* Assign individuals with highest groupness in each team to be visible (except to individual herself, who sees next highest groupness individual's contribution).

Collective actions suffer from free rider problem since not all actions are observable by others. Hence certain types of individual tend to free ride on other by exerting less effort. To prevent this situation we assign particular high-group individuals in positions/tasks where the action taken by individual is observable to others. This transparent role would serve as role model for the rest of team. This treatment allows differentiating roles within the team.

Therefore, we rank individuals by group score within the team and the contribution of highest-group individual will be displayed to others in period 1-5. The person with highest group scores observes the contribution of the second highest person within the group. In the next five periods starting from period 6, everyone observes the contribution of the two highest-group scored subjects. The highest-group scored person does see the second and third highest-group ranked subjects' contributions. The second high-group ranked person observes the first highest groupness and third highest groupness individuals contributions.

#### **Collective compensation - partial vision - with sorting (PPA) (3-2)**

---

<sup>5</sup> All sessions are held with 20 subjects except session 3 that had 16 subjects.

*Game Setup:* Shuffle VCM with punishment, 10 periods, teams of 4. Two high-gridness individuals can see contributions of all others and able to punish, rest cannot see anything in periods 1-6. Two low-gridness individuals can see contributions of all others and may punish, rest cannot see anything in periods 7-10. Total payoff cannot be seen until end.  
*Assignment Rule:* Assign individuals with highest gridness to see contributions of others.

Certain tasks and actions taken by team members in real life may turn difficult to observe. This creates a free rider problem where particular tasks can't be accomplished well. However, with some level of monitoring either random or on regular basis we could prevent slack in jobs. Therefore, we introduce a situation where a common goal within a team is achieved by combination of different roles. In particular, high-grid individuals by nature follow the norms established within a team and tend to require similar response from others. Thus, high-grid individuals inclined to monitor others and engage in inspection activity on regular basis while low-grid individuals do not. Hence, high-grid individuals at their own cost reveal unobservable actions and make it available to others. So, we assign high-grid individuals to roles where the actions of others can be seen. This way we strengthen the position of high-grid individuals who take care about social norms and withdraw from others information on free riding behavior. These way only high-grid ones are specialized in monitoring since others can't observe others actions. This treatment allows us to see whether high-grid ones really are enforcers and not. Also, we eliminate the selfish-purpose punishment, where in order to benefit in future period individuals other than high-grid may punish others. We inform subjects that there are different roles assigned to everyone and some of roles could observe actions of others.

### **Coordinative compensation with sorting 4-person (4AGA) (3-3)**

#### **(3-3a)**

*Game Setup:* N-person Assurance Game, 1 period, teams of four. Choice A (norm): 5 if everyone chooses A, or 0 otherwise. Choice B (non-norm): 3 regardless of what other players choose.

*Assignment Rule:* Concentrate high-grid individuals with one another

#### **(3-3b)**

*Game Setup:* N-person Assurance Game, 1 period, teams of four. Choice A (norm): 5 if everyone chooses A, or 0 otherwise. Choice B (non-norm): (3-number of people chosen (A)) regardless of what other players choose

*Assignment Rule:* Concentrate high-grid individuals with one another

### **Coordinative Compensation with sorting 2-person (2AGA) (3-4)**

#### **(3-4a)**

*Game Setup:* N-person Assurance Game, 1 period, teams of two. Choice A (norm): 5 if both choose A, or 0 otherwise. Choice B (non-norm): 4 regardless of what other player chooses

*Assignment Rule:* Concentrate high-grid individuals with one another

#### **(3-4b)**

*Game Setup:* N-person Assurance Game, 1 period, teams of two. Choice A (norm): 5 if both choose A, or 0 otherwise. Choice B (non-norm): 4 if both players choose (B) and 2 if other player chooses (A).

*Assignment Rule:* Concentrate high-grid individuals with one another.

An assurance game is a coordination situation where two or more players are better off if they coordinate on their decisions. There are two equilibria in this game: one is a Pareto dominant that produces higher payoff for each individual if it is enforceable for all players (everyone chooses A). There is also a risk dominant equilibrium that generates lower payoff but can provide a higher payoff at an individual level if some or all of the others follow it (choosing B). So no matter what the other players have chosen, you ensure a non-risky payoff for yourself (but lower than the Pareto optimal payoff) by choosing the second available choice rather than first available choice. On the other hand, you can maximize your payoff by all choosing A.

For the experimental condition, high-grid individuals will be clustered with other high-grid individuals, while in the control condition, individuals will be assigned at random. Participants will be divided into groups of four (3-3a, 3-3b) or two (3-3b, 3-4b). Both the grouping and the role assignment will be anonymous, meaning that no one will know which of the other people they are playing with at any time, and no one will know any other persons role at any time. In the game, each player separately and independently choose one of two options: A or B. All players will make their choices at the same time without knowing the others choice. Players will accumulate points based on the combined choices that Players make. At the end of the experiment, the points that each player accumulates will be converted to money at the rate of (1) dollars per point. Points are awarded according to the payoff structure described above under the four different treatments.

The main differences between the treatments are as follows: While the choice of A is Pareto optimal for all versions of the game, it is especially risky for the four person game, where all four members of the team must choose A in order to achieve 5 rather than 0. In the two person game, both members must choose A to achieve 5.

B is a safer choice in all treatments, but how safe it is depends on team size and the existence of a penalty. For treatment 3-3a (four-person teams), the payoff is 3 no matter how what other members choose. For treatment 3-3b (four-person teams, penalty), the payoff becomes 3 minus the number of members who have chosen A. Thus, if one member chooses B and all of the other members choose A, all members receive 0 as a payoff. For treatment 3-4a (two-person teams), the payoff is 4 no matter what other members choose. For treatment 3-4b (two-person teams, penalty), the payoff is 4 if the other member chooses B, but only 2 otherwise. See Figure 5 with four normal form games.

The assignment schedule is as follows: (i) Rank all subjects according their grid score; (ii) Assign first highest four participants with high-grid score into first team, then assign the second best ranked four subjects into second team, etc. The main implications of grid and group for this treatment can be examined by looking at the implications of choosing A. For the high group individual and the no penalty condition, choosing A may greatly benefit the other members of the team, and cannot hurt them. Substituting A and B for  $b$  and  $c$  in the generic formulation, it is clear that  $\sum_j E_A(x_j(A, a) - x_j(B, a))$  is substantially positive given any plausible method for determining  $a$ . Hence one would expect that there would be a correlation between choice of A and high groupness of individuals for the no penalty condition, but not so for the penalty conditions.

The implications of grid are less straightforward to determine, since we need to determine what the underlying norm is in this case. Next to reciprocity, the one norm that has been seen by behavioral economists to be operational across the widest range on environments is inequality aversion (Bolton and Ockenfels 2000, Fehr and Schmidt 1999), in which individuals seek to minimize the differences between their own payoffs and the payoffs of others above them, even at the risk of reducing their own absolute payoff. In both the no-penalty and penalty condition, the inequality-minimizing alternative is clearly B. However, for the no-penalty condition, it is also the risk-dominant choice, hence making it attractive even to those who have low gridness. Therefore, we would expect that there would be a correlation between gridness and choice of B that is particularly strong under the penalty condition.

#### **Design 4**

##### **Collective compensation-partial visibility-without sorting (NPPR) (4-1)**

*Game Setup:* Shuffle VCM with no punishment, 10 periods, teams of 4. One random individual's contributions are visible to everyone in periods 1-5, rest not visible, specific individuals are not announced. Two random individual's contributions are visible to everyone in periods 6-10, rest not visible, specific individuals are not announced. Total payoff cannot be seen until end.

*Assignment Rule:* random

##### **Collective compensation - partial vision (PPR) (4-2)**

*Game Setup:* Shuffle VCM with punishment, 10 periods, teams of 4. Two individuals can see contributions of all others and may punish, rest cannot see anything. Total payoff cannot be seen until end.

*Assignment Rule:* random

##### **Coordinative Compensation - 4 person (4AGR) (4-3)**

###### **(4-3a)**

*Game Setup:* N-person Assurance Game, 1 period, teams of four. Choice A (norm): 5 if everyone chooses A, or 0 otherwise Choice B (non-norm): 3 regardless of what other players choose

*Assignment Rule:* random

###### **(4-3b)**

*Game Setup:* N-person Assurance Game, 1 period, teams of four. Choice A (norm): 5 if everyone chooses A, or 0 otherwise Choice B (non-norm): (3-number of people chosen (A)) regardless of what other players choose

*Assignment Rule:* random

##### **Coordinative Compensation - 2 person (2AGR) (4-4)**

###### **(4-4a)**

*Game Setup:* N-person Assurance Game, 1 period, teams of two. Choice A (norm): 5 if both choose A, or 0 otherwise. Choice B (non-norm): 4 regardless of what other player chooses

*Assignment Rule:* random assignment

#### **(4-4b)**

*Game Setup:* N-person Assurance Game, 1 period, teams of two. Choice A (norm): 5 if both choose A, or 0 otherwise Choice B (non-norm): 4 if both players choose (B) and 2 if other player chooses (A)

*Assignment Rule:* random assignment.

## **4 Results**

Our experiments have 156 observations in eight sessions. The average grid score was 0.426 and the average group score was 0.520. The highest grid score was 1 and the lowest grid score was 0.096. The standard deviation of the grid score was 0.178. The highest group score was 0.869 and the lowest group score was 0.141. The standard deviation of group score was 0.124. Overall, the average grid score was lower than the group score and the deviation of grid score was larger than that of group score. It is important to point that these observations are the same as the previous survey and the results of World Value Survey. This result also support our earlier finding that the Grid dimension is more significant factor to differentiate cultures than the Group. We analyze data for each treatment and each design, mainly checking if the actual values coincide with the predicted values.

### **4.1 VCM Results**

#### **4.1.1 High-Group Individuals Cluster Treatment vs. Random Treatment (1-1 vs. 2-1), NP-VCM**

*Result 1 The prescription of the role-assignment algorithm for clustering individuals based on their groupness characteristics were strongly supported by the results of this treatment.*

#### **Support**

In 1-1 a sorting algorithm placed participants in groups of four based on their Group scores. The highest four group scores are being put together in one group and the next highest four together in the next group and so forth. Group one and group six represent those respondents with the highest group scores in their experimental period and group two and seven represent the second highest and so forth. Under these conditions we expected the group composed of those participants with the highest group scores (groups 1 and 6) to make the highest contributions. The overall average contributions made by groups one through ten were 30.55 and the variance between the groups was statistically significant at 1 percent level (table 3).

In 2-1 the conditions were similar to 1-1 but the groups of four were this time assigned randomly rather than by their group scores. Under these conditions we expected the overall scores for 2-1 to be lower than when we group participants based on their group scores. The average contributions for all ten groups were 25.99 and lower than the average contributions found in 1-1 (30.55). The difference in mean contributions over all ten periods was statistically significant at  $\alpha = .01$  level (based on independent means tests, see table 3). These findings strongly support our hypothesis that controlling the grouping by group scores will increase overall contributions for the entire population.

The variance between the randomly assigned groups in 2-1 was statistically significant at 1 percent level.

In both 1-1 and 2-1 treatments, mean contributions were above the Nash equilibrium level of zero. The average contributions across ten groups that had sorting algorithm (30.55) is 61 percent of the endowment and it was significantly higher than the average contribution of ten random groups in control treatment (25.99), which is only 52 percent of endowment. The Wilcoxon-Mann-Whitney rank-sum test reveals that individual contributions and efficiency with sorting in experimental group were higher than by period average contributions with no sorting in control treatment (p-value is 0.0343, two-sided in table 4). Also see Figure 1a for dynamics. This result confirms our prediction that the team with higher groupness individuals will contribute more and perform better than the team with low groupness individuals when the task requires group work and interdependence of each member on other team member's work.

#### **4.1.2 High-Grid Individuals Cluster with Low-Group Individuals vs. Random Treatment (1-2 vs. 2-2), P-VCM**

*Result 2 Over the periods in both sorting and non-sorting treatments average contributions stayed at high level relative to no punishment conditions, though clustering of high-gridness types with low-groupness types generated lower mean contributions and efficiency than the random assignment groups mean contributions.*

##### **Support**

In 1-2 the conditions were such that individuals with the low group scores were put in groups with those who had high grid scores. Two individuals with the lowest group scores were put with two individuals with the highest grid scores and then the next two highest grid scored individuals with the next two lowest group scored individuals and so forth. The assumption here was that by doing so, individuals with the highest grid scores would be put into those roles that had the greatest punishment role (since low groupness individuals would be less likely to contribute), thus increasing their positive effect on total session performance.

Groups one and six in table 2 represent those groups with two respondents with the highest grid scores and two respondents with the lowest group scores. This experiment also had a punishment component, which we expected to increase contributions over the ten periods. Under these conditions the average contribution over ten periods was 40.04. When average contributions were examined over the 10 groups the variance between the groups was statistically significant at the 1 percent level.

In 2-2 the groups were sorted randomly. Under these conditions the average contribution for all the groups was 46.00, which, contrary to our hypothesis, was higher than the average group contributions for those respondents in experimental session 1-2 (40.04 compared to 46.00 and statistically significant at 5 percent level, see table 3). The Wilcoxon rank-sum test shows that two groups averages differ at 5 percent significance level (p-value in two-sided test is 0.0178) or better in periods averages data (p=0.0002). The average contributions by group in 2-2 reflect the random nature by which the groups were assigned in that there is little in the way of a pattern although there was significant variance between all ten groups.

When examining average total payoffs for experimental design 1-2, where individuals with the lowest group scores were put in groups with those who had the highest grid scores, the highest payoffs came from the fifth group in both rounds (group 5 - 500.73 and 10 - 546.79) who were those participants with the highest group and lowest grid scores. The total average for all groups was 378.47. There was statistically significant variance between the 10 groups at the 1 percent level. Unlike the differences in total session contributions, this was not contrary to expectations, since the primary determinant of contributions is behavior in the first round, where the presence of high-group individuals in high-numbered groups would more than offset the salutary effect of high-grid individuals in low-numbered groups.

As with contributions, despite the random grouping, average total payoffs for 2-2 were actually higher than those in 1-2 (437.54 compared to 378.47, see table 3). These differences in average total payoffs were significant and were contrary to our hypothesis in that with a punishment treatment and clustering participants with high grid scores with those with low group scores, the average contributions are similar and even lower than when we randomly assign groups.

With punishment opportunity, cooperation in voluntary contribution mechanisms is much higher than without institutions. If we compare two control groups with random matching with and without punishment average contribution, efficiency and payoffs are higher with punishment than without (Wilcoxon-Mann-Whitney test  $p=0.0009$ , table 3). In line with previous findings we have higher cooperation due to the penalty threat, so that everyone is forced to contribute more when high-grid individuals are willing to punish at his/her own cost. Average contributions in sorting with punishment were also significantly higher than mean contributions in sorting no punishment treatment. However, due to confounding effect in two different algorithms in sorting conditions we did not report p-values.

Hence while the results of the clustering in the NP-VCM (1-1 vs. 2-1) provided strong confirmation of the expected significantly higher overall performance of the experimental group, assigned by an implementation of the role assignment algorithm, over the control group, assigned at random, the results of clustering in the P-VCM (2-1 vs. 2-2) were contrary to expectation. It is important to examine the possible reasons for this anomalous result. One possible explanation is the way in which the algorithm was implemented for the P-VCM did not take into account the correlations between grid and group in individuals. As previous analysis shows, there is indeed a significant (though nowhere close to unity) correlation between grid and group in individuals. Therefore, by putting high-gridness individuals with low-groupness individuals, the tendency would also be to place them with low-gridness individuals. This in turn meant that while there were more opportunities to punish low contributors, there was also less probability that these punishments would have any effect on future contributions by the punished. When we take  $x_j(b, a) - x_j(c, a)$  in the operationalization of the role-assignment algorithm, there remains a substantial amount of interpretation, particular to the nature of  $a$ . The original prediction holds only if we can assume that  $a$  itself is not endogenously affected by the nature of the assignment algorithm. In this case, there are a priori reasons that this should have been taken into consideration, which can explain the seemingly anomalous results. Attempts to begin to take this into account are provided in Section 5.

Perhaps an even bigger determinant of the results of this particular test was the specific nature of the clustering algorithm. Given the general interpretation that matching high gridness individuals with low groupness individuals, there are multiple ways in which this kind of matching can be carried out. The method chosen was to take the lowest groupness individuals first, placing them in the low-numbered teams, then take the highest gridness individuals among those remaining, then placing them in the same teams. This essentially magnifies any correlation between gridness and groupness, thus making it possible that the low-numbered teams could actually have similar or even lower average gridness than the high numbered teams, while their average groupness would clearly be lower. Under such conditions, the predictions for overall performance across the session would be confounded. What this outcome points is the importance of how the role-assignment algorithm is implemented, and that the operationalization rules for the algorithm need to be corrected and extended to deal with the level of complexity brought upon by clustering when individuals characteristics are not independently distributed.

We know that high-grid scored individuals can be singled out through the punishment task in the shuffled treatments from the self-interested low-gridness motive to punish others. In those shuffled treatments the rate of punishment was low compared to partner treatments, therefore cooperation was lower with shuffled compared to partner treatments. So this algorithm that sorts high-gridness subjects with low-groupness subjects was implemented in partner treatments, hence did not exclude selfish motive to punish others. Overall in the sorted groups punishment cost was much higher than in the unsorted groups (Wilcoxon rank-sum test  $p=0.0019$ ). The regression coefficient on "positive deviation" from others average was significant and negative in the partners treatment while in the shuffled conditions the coefficient was positive and insignificant. This suggests that there was a systematic punishment toward high contributors in the partner treatments that could remove punishment effect. Wilcoxon-Mann Whitney test shows that in sorted groups where high gridness were clustered with low-groupness, punishing cost incurred by low-groupness-low-gridness (selfish) subjects was much higher than high-gridness (strong reciprocators) subject's punishing cost ( $p=0.0004$ ). The lowest punishing cost was exhibited by low-grid-high-groupness (altruists) subjects (comparing mean punishing cost of altruist and strong reciprocator,  $p=0.0005$ ). This again demonstrates that antisocial punishment was in place. In particular, 9.3 percent of actions were punished despite the contribution level was equal to the full endowment in the sorted partner groups while 5.3 percent of high-contributors were punished in the shuffled sorted groups. These numbers were 5.75 and 7 percent in the random partner and shuffled groups respectively. It seems in design 2 with punishment condition antisocial punishment drove down the payoffs and efficiency.

The average profits in the random groups were higher than in the sorted groups before the punishment stage ( $d_3 < d_4 < d_1 < d_2$ , Wilcoxon sign ranks test  $p=0.0051$  for all pair wise comparisons where  $d_1$  indicates per period average profits before the punishment stage in the partner sorted groups (design 1)). After the punishment stage and payoff reductions the mean profits in the control groups were still higher than in the partner matching sorted groups while in the shuffled treatments there was no difference in profits between sorted and unsorted groups ( $d_3 = d_4 = d_1 < d_2$ , Wilcoxon signed ranks test  $p$ -values for pairwise comparisons are in the range [ 0.0007, 0.0015 ] ). This can be

explained by very harsh punishment done in the sorted unchanged groups such that their ranking moved from the second position to the last rank and fell down to the shuffled groups level.

### **High-Group Individuals Visible Treatment vs. Random Treatment (3-1 vs. 4-1), NP-VCM**

*Result 3 Visibility of only high-groupness individual's contributions within the team produced similar performance with random-visibility assignment.*

#### Support

In 3-1 there was no punishment over the ten periods of the VCM. What differentiated this from a traditional NP-VCM (where all individuals' contributions are visible to everyone) was that only one individual's contribution was visible to everyone over the first five periods and only two individual contributions were visible over the last five periods. Individuals with the highest groupness scores on each team were the ones would were made visible. For periods 1-5, it was the member of the team with the single highest groupness score. For periods 6-10, it was the two members with the two highest scores. The clustering into teams was assigned randomly.

The idea behind this role assignment is that placing high-groupness individuals in positions of high visibility would put them in a position to trigger existing gridness-based reactions among other members of the team to meet the norm, and hence would increase overall contribution. Here, however, the effect of the visible role  $i$  on  $x_j(b, a) - x_j(c, a)$  is not direct, but rather due to its endogenous effect on  $\mathbf{a}$  given non-zero gridness among other team members. Under these conditions, the average contribution for all the groups was 27.8. When examining the contributions by group, the contributions varied statistically significantly at 1 percent level.

Like 3-1, in 4-1 there was again no punishment over the ten periods. The difference was that the one individual whose contribution was visible during periods 1-5 was chosen randomly. Likewise, two individuals' contributions were randomly chosen to be visible to the rest of the group during periods 6-10. The groups of four were randomly assigned. Under these conditions the average contributions overall was 24.6. Therefore the average contribution in 4-1 was lower than the average contribution 3-1 (27.8 compared to 26.4, see table 3). The variance within the groups was statistically significant at 5 percent level.

Wilcoxon rank-sum test shows that the average contributions across ten sorted groups was no different than the average contribution of ten random groups in control treatment ( $p$ -value=0.9397, table 4). The efficiency levels were no different across sorted and unsorted groups. This result has no contrary to our hypothesis that putting high-groupness individuals in roles that have greater visibility will trigger higher cooperation among others than low-group individuals will. Also, it consistent with the idea that changes in information on past contribution will affect contribution at a rate that is positively related

to gridness. One reason of no visibility effect may relate to the fact that the grid scores<sup>6</sup> in the control groups were significantly higher than the grid scores in the sorted groups (table 7,  $p=0.0243$ ). In design 3 and 4 higher grid score may have created more responsiveness in control groups, therefore sorting had no effect. This suggests difficulty of controlling natural characteristics that subjects bring with them into lab. To verify this we check if contributions of high-gridness subjects increase over time in the unsorted groups. Hence, with less information on low-group individuals contributions and visibility of highest group score individuals' actions, conditional cooperators were not able to reciprocate others and the decay could be prevented by sorting algorithm while random visibility condition increases decay. With partial visibility average contributions reached 27.8 (56 percent of endowment) while with full visibility in the previous study (Chai et al. 2008), the mean contributions in the shuffled treatments was 21.5 (43 percent of endowment). In partner treatments with full information cooperation was about at 30 (61 percent of endowment).

#### **4.1.3 Two High/Low-Grid Individuals Vision vs. Two Random Vision Treatment (3-2 vs. 4-2), P-VCM**

*Result 4 In line with predictions, control group outperforms the experimental group when low-gridness individuals are assigned to punish others. Contrary, assignment of high-gridness individuals into norm enforcer's role produces lower contributions than the random schedule. However, the efficiency in the sorted and unsorted groups were no different. As predicted, with punishment opportunity output is higher than without.*

#### **Support**

In treatment 3-2, periods 1 through 6, only those two individuals with the highest grid scores on could see the contributions of all other members on a team and were able to punish them. Those two individuals with the lowest grid scores could neither see others contributions nor punish them. In periods 7-10, the tables were turned: the two individuals with the lowest grid scores could see the contributions of all other members and could punish participants, while those with the highest could not see or punish. In experiment 4-2, two random individuals could see the contributions of all other members and could punish them, while the rest could not see anything nor punish. It was expected that average contributions would be higher in the experimental treatment than the control in periods 1 through 6, since high-grid individuals would be placed in a position to sanction other individuals, and for similar reasons average contributions would be lower in the experimental treatment in periods 7-10. Because of the tendency of contributions to rise over time in a P-VCM from periods 1-10, no attempt was made to compare the contributions of periods 1-6 and periods 7-10. The membership of teams was assigned randomly for both experimental and control treatments.

---

<sup>6</sup> In design 1 and 2, grid/group scores in experimental groups vs. control groups were no much different from each other (Wilcoxon rank sum one-sided test p-values are 0.528 and 0.593 respectively for grid and group scores). Only grid scores in design 4 (sessions 4 and 8) were significantly higher than the grid scores for the design 3 (sessions 2 and 5).

Average contributions over the first 6 periods in treatment 3-2 (see table 4), where individuals with the highest grid scores could see contributions and punish other participants, were actually lower than the average contributions over the first six periods in session 4-2 (35.49 and 38.94 respectively). These results are in the opposite of our hypothesized direction. We expected vision and ability to punish among high grid participants to increase the contributions for the entire group, yet the control group contributed more.

Average contributions over the final four periods in treatment 3-2, where two participants with low grid scores could see all contributions and punish, were significantly different from average contributions over the final four periods in session 4-2, where contribution visibility was done randomly (38.07 compared to 40.81). As expected, Wilcoxon rank-sum test shows that control group in last four rounds did perform better than experimental group with low-grid assignment at all significance levels (p-value<0.0000, one-sided).

As with the results in 1-2 and 2-2, the anomalies may be placed at the endogeneity of actions by those who were not in a position to see contributions or punish. For treatment 3-2, periods 1-6, the effect is even more straightforward and should have been foreseen: those who were not given vision had the lowest gridness scores in their teams. Hence, while they were in a position to be punished and more likely to receive penalties from the high-grid members if they failed to contribute, they were also the least likely to respond to those penalties by raising their contributions. Hence, once again, the need to take into account the endogeneity of a into the operationalization of the role assignment algorithm was highlighted. Also, as mentioned before the grid scores in control sessions were happen to be significantly stronger than in experimental groups which lead to more punishment and more responsiveness from punished side in control sessions. In addition, threat of punishment induces everyone to behave according the norms which results in higher contributions in both control and experimental groups. The punishment costs were no different across sorted and unsorted groups.

With punishment opportunity performance is generally expected to be much higher than without punishment institutions. Thus, random control group with punishment performs better than the random control group with no institutions. Wilcoxon-Mann-Whitney two-sided test is highly significant with p-value=0.0002.

## 4.2 Individual behavior

We identify cultural types and classify them based on two criteria: (i) answers provided during the pre-treatment survey; (ii) actions taken by subjects within each VCM treatments. By our definition, cooperative action refers to the contribution amounting more than fifty percent of endowment. Contribution which is less or equal to 25 tokens will be referred to the non-cooperative action. Our cut-point of fifty percent of endowment is similar to Houser and Kurzban (2005)<sup>7</sup>. In addition classification of behavioral types incorporates institutional feature as follows. Unconditional cooperators

---

<sup>7</sup> See also Gunnthorsdottir et al. (2007) where cut-point was 30 or 33 percent of endowment (Isaak and Walker 1988), or other classifications in different context El-Gamal and Grether, 1995, or Casari and Plott (2003) for spite, altruist and selfish motives

are the ones who do cooperate in more than fifty percent of time regardless of institution imposed. Free riders do not cooperate in both treatments in more than fifty percent of time. Conditional cooperators or profit maximizers cooperate in P-treatment and do not cooperate in NP-treatment. Irrational subjects cooperate with no punishment and do not with punishment.

*Result 5 Cooperation changes with institutions: in regular VCM contributions decline toward the end of rounds while in VCM with punishment subjects sustain cooperation until the end of game. This difference in behavior can be explained by interaction of cultural types. High-group subjects contribute more and high-grid subjects punish more than other types.*

Support

*Contribution:* Figure 1a, 1b. Individual actions were diverse and vary from zero to full endowment contribution. In our sample 51.9 percent were cooperators, 14.1 percent had non-cooperative behavior, 33.3 percent were reciprocators that condition cooperation on situation, and 0.6 percent of population had irrational behavior. Our results are robust if we change the definition by picking up a single period or average action. To make a comparison, in our sessions 22.4 percent of population had low-grid and high-group, 27.6 percent had low-grid and low-group, 21.2 percent had high-grid and low-group, and 28.8 percent had high-grid and high-group characteristics. Latter classification based on the cut-point equal to the average grid and group scores for the whole samples that were 0.426 and 0.52 respectively<sup>8</sup>.

With punishment opportunity 76 percent of all actions were contributions above 35 tokens compared to 43 percent in no punishment treatment. Also percent of full free riding actions ( $c_i = 0$ ) dropped from 17 percent in no punishment VCM to 3.6 percent with punishment. Reciprocators are the ones who accomplish this major drop since full free riding actions of high-grid subjects regardless of groupness score fall from 7.2 percent to 1.9 percent with punishment. If we look at the classification based on actions taken during experiments percentage drop in full free riding by reciprocators are pretty much the same respectively, 11 and 0.7. This verifies again the predictive power of survey instrument. Overall in the pooled data for no punishment treatments Wilcoxon signed ranks test shows that high-groupness individuals exhibit significantly higher level of contributions than low-groupness individuals ( $p=0.0469$ , two-tailed)

*Punishment:* Figure 4, Table 6. Ordinary least square (OLS) regression analysis shows negative relationship between punishment points that subjects received from others and negative deviation from other's average contribution. See Table 6 and Figure 4. We run ordinary least square (OLS) regression of received points on constant, other's average contribution, absolute negative deviation, positive deviation and dummies for rounds and sessions in shuffled treatments, dummies for rounds and matched groups in partner treatments. In both shuffled and partner treatments the coefficient of the absolute negative deviation is highly significant and positive indicating that subjects were punished more the more their contribution fall below the other's average. In both shuffled

---

<sup>8</sup> We consider grid/group score to be high if subject's score calculated using grid/group indexes are above the average grid/group score of the sample and low if it's below that average score.

and unchanged groups there was significant impact of ‘‘others’ average contributions’’ on received punishment points; moreover negative coefficient indicates an attempt to establish higher norms. The coefficient on positive deviation is insignificant and low in shuffled treatment which confirms absence of any systematic antisocial, spiteful or revenge-type punishment. However, in partner treatments there is a systematic evidence of punishing less those who contribute more. After-treatment dominant response on motives to punish others was to punish those who were selfish, not investing in group exchange and contributing way less than the average.

As predicted average punishing cost incurred by high-gridness subjects was significantly higher than punishment performed by low-grid-high-group individuals in pooled data (Wilcoxon signed-rank test  $p$ -value=0.0069). Low-grid- low-group subjects punished others more than low-grid-high-group subjects as well (Wilcoxon signed-rank test  $p$ -value=0.0051). Interestingly, the punishing cost incurred by high-gridness subjects was lower than low-grid-low-group individual’s punishing cost (Wilcoxon signed-rank test  $p$ -value=0.0166). This suggests that low-grid-low-group individuals in addition to high-gridness individuals can guard social norms and orders. Median test shows significant difference in punishing behavior among three cultural types distinguished by grid/group scores ( $p$ -value for the Pearson chi-squared test=0.006). Here data is pooled , therefore we have selfish motive to punish others in the partner matching. In the shuffled groups punishing cost incurred by high-gridness was again lower from low-grid-low-group type (Wilcoxon rank-sum test  $p$ =0.0031). In the partner treatments above costs were no different.

## 4.5 Coordinative Compensation Game Results

This section reports assurance experiments where participants in groups of four or two have experienced four different assurance games with the normal form structures depicted in table 10. First, they experience 4-person games, next - 2-person games in the same order of the normal forms depicted in Figure 5.

*Result 8 High groupness subjects select Pareto optimal payoff choice than low groupness ones. High gridness individuals choose non-risky payoff choice. The size of the group negatively affects coordination on Pareto optimal outcome.*

### 4.5.1 Support

*Grid/group and coordination:* (Table 8, 9and 10). There were 80 subjects who joined a coordinative compensation game in four sessions. The average grid score of these subjects was 0.431 and the standard deviation was 0.171. The average group score was 0.531 and the standard deviation was 0.132. The highest and the lowest grid score were 1 and 0.095, respectively. The highest and the lowest group score were 0.869 and 0.273, respectively.

In the assurance game choice ‘‘A’’ generated Pareto optimal payoff while choice ‘‘B’’ produced risk-dominant payoff. In the 4-person game with no penalty for picking ‘‘B’’ when others failed to do so, 63 percent out of 80 subjects chose ‘‘A’’ and the average

payoff was 2.40 dollars. In the game with penalty, the number of people who chose "A" increased to 83 percent and the average payoff increased to 2.50 dollars. In the 2-person game, when the return of choosing "A" was higher, 44 percent out of the 80 subjects chose "A" and the average payoff was 2.29 dollars. In the game with penalty, the number of people who chose "A" increased to 51 (64 percent) and the average payoff increased to 3.30 dollars.

Generally, 2-person game had higher average payoff (5.90) than the 4-person game (5.10), which is not surprising because the smaller group, the higher possibility to reach the matching position.

Considered together with gridness characteristics, as Table 8 shows, the correlation between the choice "B" and the grid score was positive, and the correlation for the 4-person game with penalty and pooled data were significant at the 5 percent level, which means, high-grid score individuals tend to choose "B". Moreover, the significance was achieved under the penalty condition, as well as the pooled data. Both were in accord with the predictions of our implementation of grid and group. The fact that significance was achieved under the 4-person group follows intuitively from the logic of the model, since under the 2-person group, there was a relatively larger chance of both achieving equality with the fellow team member and a higher absolute payoff by picking "A".

The results also reflect that groupness has a statistically significant relationship with subject's choice. As shown in the following Table 9, the coefficients indicate that high-group individuals would also be likely to pick "A". This result was more significant in the game where there was no penalty to others, since the choice of "A" will never hurt and sometimes benefits the other group members. Although none of the coefficients reach the  $p < 0.05$  level of significance, p-value for the no penalty, 2-person game was 0.0504. This is also in accordance with our expectations. The pooled effect is significant for the 2-person group, which makes sense, since the conditions under which the partner benefits make up a greater portion of possible outcomes (one out of two instead of one out of eight).

## **5 Implications for Grid-Group Formalization, Role-Assignment, and Experimental Design**

The results of these experiments provided strong endorsement for the current approach to the choice-theoretic formalization and measurement of grid and group. In all the experiments where the indexed and alternative approaches to measuring grid and group were compared, their implications for behavior were similar, demonstrating the robustness of the two concepts as they are being used. While this does not rule out continued testing and elaboration of both measurement and formalization, there is no clear evidence that the current approach needs to be changed before proceeding with future empirical research.

The role-assignment algorithm as a whole performed as anticipated. Operationalized to generate recommendations for intra-team and inter-team assignments, it general tended recommend assignments that increased team performance. The concepts of dependence and coordination continue to be robust ones for representing

characteristics of roles that make them suited to high-groupness and high-gridness individuals, but further work needs to be done in the operationalization of these concepts both for formal experimental conditions and for less precisely defined real-world conditions. In particular, the measures must more explicitly take into account the extent to which the relationship between a roles actions and team performance depends not only on the institutional and structural constraints impinging on the actor, but also on calculation of the actions of the other members of the team. In other words, while the framework is already incorporated into a game-theoretic formalization, it must use delve more deeply into the technology of equilibrium determination and best response determination in order to make more precise predictions.

Likewise, this means taking into consideration the balance between individual and team incentives when examining the role of groupness. A better suitability of role to high groupness could be generated by looking at the extent to which a roles actions that best facilitate team performance are contrary to those that would be optimal for the roles individual payoff.

Likewise, the way in which coordination relates to gridness could be modified. In the PVCN experiments, we examined the importance of gridness in helping to perpetuate the norm of reciprocity, even in cases where it is not in the interest of an individual to follow the norm. This in turn promotes cooperative behavior from others and hence higher performance from the team. However, this goes beyond coordination in the conventional sense and takes into account the common knowledge of rationality between members of team. This could be determined looking at the effect on team performance of the actions that would become rational to other team members conditional on the action of the role. Both the public vs. private incentive issue and taking into account the effects of a roles actions on other rational team members apply to both grid and group. From the operationalization point of view, this would mean elaborating the definitions so that  $v_{ij} = E_A((x_j(b, a|b)) - x_j(c, a|c)) / (x_i(c, a|c) - x_i(b, a|b))$  and  $\mu_{ij} = Var_A((x_j(b, a|b)) - x_j(c, a|c)) / (x_i(c, a|c) - x_i(b, a|b))$ , where  $(a|b)$  is a conditional probability distribution over actions by the other members of the group, given that the individual in the designated role has performed  $b$ , and similarly for  $(a|c)$ . Of course, determining a conditional probability distribution assumes the existence of priors, and it is often not enough to know the structural conditions surrounding a team to calculate what those priors are here we may need to go beyond the realm of grid and group (as well as these have served us) and examine specific the specific ideological milieu surrounding an organization. This in turn will mean not only complicating the algorithm, but an active engagement with qualitative information and subject matter experts who can provide insights into these ideologies.

One specific problem that arises in implementing a role-assignment algorithm is how to sequence selection for particular roles. Sequencing problems, along with the effects of correlation between grid and group, may have contributed to the two major anomalies of the Stage II studies, the clustering of high grid with low group in 1-2 and 2-2 and the selective vision provided to high grid team members in 3-2 and 4-2. Better sequencing means paying attention to the effects of a role selection on the pool of individuals left for the remaining roles, and the effects this will have on the kinds of interactive responses that the selected role will face from fellow team members. There

is unlikely to be any general optimal methodology for sequencing, which in 50 turn makes it more difficult to determine the general optimal solution to the problem of role selection analytically. This may necessitate the use of heuristics for locating local optima and a process that contains elements of satisficing rather than pure optimization in role assignment.

Finally, improvements can be made in the design of the experiments themselves. In order to conserve time and money resources, multiple treatments were conducted within a single session during this project. In particular, with the Stage II experiments, the fact that subjects were confronted with multiple treatments with complicated instructions may have led to confusion or exhaustion that could account for the presence of outlier behaviors and the general lack of significance. Although efforts were made to follow treatments of a particular kind (NP-VCM, P-VCM, UG, Trust, Assurance) with different kind in order to minimize any tendency to carry over strategies from one treatment to another, this meant that subjects had to understand a wider range of treatment conditions. With greater resources, treatments could be split up more finely into separate sessions while maintaining sufficiently large sample size. Of course, even with greater resources, a tradeoff between implementation efficiency and possible validity issues will exist at some level. This in turn may be addressed by further work on experimental designs that simplifying the amount of information each role requires while maintaining the desired separation between expected behaviors vis-a-vis cultural differences.

#### Reference

- [1] Andreoni, J., 1988. Why Free Ride? Strategies and Learning in Public Goods Experiments. *Journal of Public Economics* 37, 291-304.
- [2] Andreoni, J., 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? *The American Economic Review* 85, 891-904.
- [3] Andreoni, James, Marco Castillo, Ragan Petrie, 2003. What Do Bargainer's Preferences Look Like? Experiments with a Convex Ultimatum Game. *American Economic Review* 93 (3), 672-685.
- [4] Ashraf, Nava, Iris Bohnet and Nikita Piankov. 2006. Decomposing Trust and Trustworthiness. *Experimental Economics* 9, 193-208.
- [5] Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122-142.
- [6] Binmore, Ken. 2007. *Does Game Theory Work? The Bargaining Challenge*. Cambridge, Mass.: The MIT Press.
- [7] Bohnet, I., Kuebler, D., 2005. Compensating the Cooperators: Is Sorting in the Prisoner's Dilemma Possible? *Journal of Economic Behavior and Organization*, 56: 61-76.
- [8] Bolton G. 1991. Comparative model of bargaining: Theory and evidence. *American Economic Review* 81, 1096-1136.
- [9] Bolton, Gary E. and Axel Ockenfels, 2000. ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, American Economic Association, 90(1/March), 166-193.

- [10] Buchan, N., Johnson, E., Croson, R, 2006. Getting to know you: an international experiment of the influence of culture, communication, and social distance on trust and trustworthiness. *Journal of Economic Behavior and Organization*, 60: 373-398.
- [11] Carpenter, Jeffrey, Samuel Bowles, and Herbert Gintis, 2007. Mutual Monitoring in Teams: Theory and Experimental Evidence on the Importance of Reciprocity, IZA Discussion Papers 2106, Institute for the Study of Labor (IZA)
- [12] Carter J. R., and Irons M. D., 1991. Are economists different, and if so , why?. *Journal of Economics Perspectives* 5, 171-177.
- [13] Chai, Sun-Ki and Aaron Wildavsky, 1994. Culture, Rationality and Violence In Dennis J. Coyle and Richard J. Ellis (eds.), *Politics, Culture and Policy: Applications of Cultural Theory* (Boulder, CO: Westview Press, 1994). Reprinted in Chai and Swedlow (ed.), *Culture and Social Theory*, 281-98.
- [14] Chai, Sun-Ki and Brendon Swedlow (eds.), 1998. *Culture and Social Theory*. New Brunswick, NJ: Transaction Publishers, 1998.
- [15] Chen, X. P. 1996. The group-based binding pledges a solution to public goods problems. *Organizational Behavior and Human Decision Processes* 66, 192-202.
- [16] Douglas, Mary, 1970. *Natural Symbols: Explorations in Cosmology*. New York: Pantheon
- [17] Douglas, Mary, 1980. Introduction to Maurice Halbwachs *The Collective Memory*, Harper and Row.
- [18] Douglas, Mary and Aaron Wildavsky, 1982. *Risk and Culture: An Essay on the Selection of Technical and Environmental Dangers*. Berkeley: University of California Press.
- [19] Dufwenberg, Martin and , Georg Kirchsteiger. 1998. A Theory of Sequential Reciprocity, CentER Discussion paper No. 9837, Tilburg University
- [20] Dufwenberg, Martin and Uri Gneezy. 2000. Measuring Beliefs in an Experimental Lost Wallet Game. *Games and Economic Behavior*, 30:163- 182.
- [21] Eckel, C., Grossman, P., 1998. Are women less selfish than men? Evidence from dictator experiments. *Economic Journal* 108, 726-735.
- [22] Falk, Armin and Fischbacher, Urs. 1998. A theory of reciprocity. Institute of Empirical Research in Economics. University of Zurich. Working Paper No.6.
- [23] Fehr, Ernst, Urs Fischbacher and Simon Gächter (2002) Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms, *Human Nature* 13(1), 1-25.
- [24] Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. *American Economic Review*, 90 (4): 980-994.
- [25] Fehr, Ernst and Gintis, Herbert, 2007. Human Motivation and Social Cooperation: Experimental. and Analytical Foundations, *Annual Review of Sociology*, 33: 43-64.
- [26] Fehr, E., K. Schmidt, 1999. A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114(3), 817-851.
- [27] Fischbacher, Urs and Simon Gächter 2006. Heterogeneous social preferences and the dynamics of free riding in public goods Institute for Empirical Research in Economics, University of Zurich, Working Paper Series ISSN 1424-0459 Working Paper No. 261
- [28] Fischbacher, U., Gächter, S., Fehr, E., 2001. Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. *Economics Letters*, 71: 397-404.

- [29] Forsythe R., Horowitz J.L., Savin N.E., Selton M. 1994. The statistical analysis of experiments with simple bargaining games. *Games and Economic Behavior*, 6: 347-369.
- [30] Fukuyama, F., 1995. *Trust: The Social Virtues and the Creation of Prosperity*. The Free Press, New York.
- [31] 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.
- [32] Gunnthorsdottir, Anna, Houser, D., McCabe, K., and Ameden, H. 2007. Disposition, History, and Contributions in Public Goods. *Journal of Economic Behavior and Organization*, 62: 304-315.
- [33] Guth W., Schmittberger R., and Schwarze B., 1982. An Experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3, 367-388.
- [34] Guth W., and Tietz R., 1990. Ultimatum bargaining behavior: A survey as comparison of experimental results. *Journal of Economic Psychology* 11, 417-449.
- [35] Hardin, G., 1968. The Tragedy of the Commons. *Science*. 162, 1243-48.
- [36] Henrich, Hoesep; Boyd, Robert; Bowles, Samuel; Camerer, Colin; Fehr Ernst; Gintis, Herbert and Richard McElreath. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review* 91(2): 73-78.
- Hermann, Benedikt , Christian Thöni, Simon Gächter, 2008 Antisocial Punishment Across Societies. *Science*, 309: 1362-1367.
- [37] Hoffman, E., McCabe, K., Smith, V.L., 1996. Social distance and other-regarding behavior in dictator games. *American Economic Review* 86, 653-660.
- [38] Inglehart, R., Basanez, M., Moreno, A., 1998. *Human Cultures and Beliefs: A Cross-Cultural Sourcebook: Political, Religious, Sexual, and Economic Norms in 43 Societies: Findings from the 1990-1993 World Values Survey*. University of Michigan Press, Ann Arbor.
- [39] Inglehart, Ronald., Basanez, M., Diez-Medrano, J., Halman, L. and Luijkx, R. 2004. *Human Beliefs and Values: A Cross-Cultural Sourcebook based on the 1999 2002 Values Surveys*, Siglo XXI Editores: Mexico.
- [40] Kelley, Harold H. and Anthony J. Stahelski. 1970. Social Interaction Basis of Cooperators' and Competitors' Beliefs About Others. *Journal of Personality and Social Psychology*, 16:66-91.
- [41] Kreps, D. M. 1990. Corporate Culture and Economic Theory, in *Perspectives on Positive Political Economy* (J. Alt and K. Shepsle, Eds.). Cambridge: Cambridge Univ. Press
- [42] Kurzban, R., Houser, D., 2002. Revisiting Kindness and Confusion in Public Goods Experiments. *The American Economic Review* 92 (4), 1062-1069.
- [43] Kurzban, R., Kevin McCabe, Vernon L. Smith, Bart J. Wilson, 2001. Incremental Commitment and Reciprocity in a Real-Time Public Goods Game. *Personality and Social Psychology Bulletin* 27 (12): 1662-1673.
- [44] Kurzban, R., Houser, D., 2005. An Experimental Investigation of Cooperative Types in Human Groups: a Complement to Evolutionary Theory and Simulation. *Proceedings of the National Academy of Sciences* 102, 1803-1807.

- [45] Ledyard, J.O., 1995. Public goods: A survey of experimental results. In: Kagel, J.H., Roth, A.E. (Eds.). *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- [46] McCabe, K., and Smith, V., 2000. Goodwill accounting in economic exchange. In: Gigerenzer, G., Selten, R. (Eds.), *Bounded Rationality: The Adaptive Toolbox*. MIT Press, Cambridge, MA, 319-340.
- [47] Olson, Mancur. 1971 [1965] *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.
- [48] Ostrom, Elinor (ed.), 2007. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.
- [49] Prasnikar V. and Roth A. E., 1992. Considerations of fairness and strategy: Experimental data from sequential games. *Quarterly Journal of Economics* 35, 309-332.
- [50] Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83(5), 1281-1302.
- [51] Rigdon, M., McCabe, K., Smith, V. 2007. Sustaining Cooperation in Trust Games. *The Economic Journal*, 117 ( 522): 991-1007.
- [52] Rosenthal, R. 1982. Games of Perfect Information, Predatory Pricing, and the Chain-Sote Paradox, *Journal of Economic Theory* 25: 92-100.
- [53] Van Huyck, J., Battalio, R., and Walters, M. 1993. Commitment versus Discretion in the Peasant-Dictator Game: Aggregate Analysis, mimeo, Texas A and M Economics.
- [54] Yamagishi, T., and Sato, K., 1986. Motivational Bases of the public goods problem. *Journal of Personality and Social Psychology* 50, 67-73.
- [55] Yamagishi, M., and T. Yamagishi, 1988. The Provision of a Sanctioning System in the United States and Japan. *Social Psychology Quarterly* 5 (1), 265-71.
- [56] Yamagishi, T., 1988. The provision of a sanctioning system in the United States and Japan. *Social Psychology Quarterly* 51, 32-42.
- [57] Yamagishi, T., Cook, K., Watabe, M., 1998. Uncertainty, trust, and commitment formation in the United States and Japan. *American Journal of Sociology* 104, 165-194.

Table 1: Summary of Experimental Sessions

Treatments	Periods	Groups	Date	Session
Design 1:	NP, P, T, UG			
(NP-A*), (P-A*)	10,10	5	3/11/08	1
	10,10	5	3/13/08	6
Design 2:				
(NP-R*), (P-R*)	10,10	4	3/12/08	3
	10,10	5	3/14/08	7
Design 3:	NP, P, 4AG, 2AG			
(NPP-A), (PP-A), (4AG-A), (2AG-A)	10,10,2,2	5	3/11/08	2
	10,10,2,2	5	3/13/08	5
Design 4:				
(NPP-R), (PP-R), (4AG), (2AG)	10,10,2,2	5	3/12/08	4
	10,10,2,2	5	3/14/08	8

NP-no punishment, P-punishment, NPP-no punishment partial visibility, PP-punishment partial vision, A- algorithm, UG-ultimatum game, Trust-trust game; \*-partner (iterated) condition

Table 3. Average Contributions by Experiment Design & by Group (1-2 & 2-2 also includes Average Total Profit)

design groups	No Punishment		Punishment		Punishment		No Punishment		Punishment	
	1--1 NPA	2--1 NPR	1--2 PA	2--2 PR	Total Ind. Payoff PA (1-2)	PR(2-2)	3--1 NPPA	4--1 NPPR	3--2 PPA	4--2 PPR
1	28.77***	36.92***	37.88***	31.5***	340.34***	259.17***	15***	29.13*	32.55***	42.17
2	40.9***	32.25***	41.88***	48.08***	469.76***	534.92***	24.73***	21.8*	30.25***	41.6
3	41.47***	8.8***	42.53	48.5***	437.2***	506.42***	23.02***	28.9*	32.28***	41.92
4	26.25***	21.42***	30.83	46.42***	248.9***	472.06***	19.48***	15.27*	31.63***	42.1
5	32.28***	~	44.1	~	500.73***	~	26.48***	28.15*	30.23***	40.58
6	27.75***	30.65***	26.6	48.47***	-12.22***	455.49***	35.83***	28.5*	41.35***	38.33
7	25.5***	28.55***	39.75	46.42***	434.38***	502.94***	31.3***	27.42*	42***	36.88
8	36.72***	40.08***	43.13	47.5***	463.28***	460.32***	38.58***	27.45*	40.9***	36.67
9	24.77***	26.1***	43.75	48.88***	355.57***	520.59***	35.47***	28.98*	42.5***	37.88
10	21.08***	9.13***	49.97	48.25***	546.79***	543***	27.8***	28.42*	41.55***	38.78
mean	<b>30.5</b>	<b>26.0</b>	<b>40.0</b>	<b>46.0</b>	<b>378.5</b>	<b>472.8</b>	<b>27.8</b>	<b>26.4</b>	<b>36.5</b>	<b>39.7</b>
Ho:	NPA=NPR	NPA>NPR	PA=PR	PA>PR	PA=PR	PA>PR	NPPA=NPPR	NPPA>NPPR	PPA=PPR	PPA>PPR
p <sup>1</sup> -value:	0.5136	0.589	0.0178	0.1780	0.0724	0.256	0.9397	0.5100	0.2899	0.3600
Ho:		NPR=PR		NPR<PR				NPPR=PPR		NPR<PR
p <sup>1</sup> -value:		0.0009		0.9630				0.0002		1.0000

Note: <sup>1</sup> Wilcoxon-Mann-Whitney rank-sum test, two-sided and one-sided

Note: Significance refers to significant variance between groups within 1-1, 2-1, 3-1 etc. based on one-way ANOVA test

Note: \*\*\* refers to p value<0.001, \*\* refers to p value<.01, \* refers to p value<.05

NPA-no punishment with algorithm, NPR-no punishment random, PA-punishment with algorithm, PR-punishment random, NPPA-no punishment partial visibility algorithm, NPPR-no punishment partial visibility random, PPA-punishment partial vision algorithm, PPR-punishment partial vision random.

Table 4. Averages in VCM by treatment

mean/design	No Punishment		Punishment		No Punishment		Punishment	
	1--1	2--1	1--2	2--2	3--1	4--1	3--2	4--2
	NPA	NPR	PA	PR	NPPA	NPPR	PPA	PPR
Contribution	<b>30.5</b> (17.6)	<b>26.0</b> (18.8)	<b>40.0</b> (15.4)	<b>46.0</b> (9.8)	<b>27.8</b> (17.9)	<b>26.4</b> (19.2)	<b>36.5</b> (14.7)	<b>39.7</b> (13.4)
Efficiency	<b>71.0</b> (30.2)	<b>86.1</b> (20.0)	<b>75.2</b> (21.0)	<b>77.9</b> (22.0)	<b>80.6</b> (10.1)	<b>76.0</b> (13.0)	<b>77.8</b> (10.6)	<b>76.4</b> (9.8)
Cum. Payoff	<b>709.8</b> (273.4)	<b>860.7</b> (164.0)	<b>751.6</b> (92.00)	<b>779.3</b> (110.3)	<b>805.5</b> (118.4)	<b>759.9</b> (135.9)	<b>777.7</b> (95.3)	<b>764.0</b> (84.5)
Punish			<b>6.80</b> (15.6)	<b>3.4</b> (7.2)			<b>4.0</b> (14.4)	<b>4.2</b> (14.6)

**Compare mean period values**

p*-value\Ho:	NPA=NPR	NPA>NPR	PA=PR	PA>PR	NPPA=NPPR	NPPA>NPPR	PPA=PPR	PPA>PPR	
Contribution	0.0343	0.7800	0.0002	0.0000	0.2899	0.6400	0.0015	0.0800	
Efficiency	0.0343	0.7800	0.0004	0.0300	0.2899	0.6400	0.2899	0.3600	
Cum. Payoff	0.0343	0.7800	0.0004	0.0300	0.2899	0.6400	0.2899	0.3600	
Punish			0.0019	0.9100			0.6229	0.4350	
p*-value\Ho:	NPR=PR		NPR<PR		NPPR=PPR		NPPR<PPR		
Contribution	0.0002		1		0.0002		1.000		
Efficiency	0.0015		0.9200		0.2899		0.6400		
Cum. Payoff	0.0015		0.9200		0.2899		0.6400		
p*-value\Ho:							PPA=PPR	PPA>PPR	
Contribution							first 6 rounds last 4 rounds	0.0039 0.0209	0.0000 0.0000

**Compare mean group values**

p*-value\Ho:	NPA=NPR	NPA>NPR	PA=PR	PA>PR	NPPA=NPPR	NPPA>NPPR	PPA=PPR	PPA>PPR	
Contribution	0.5136	0.589	0.0178	0.178	0.9397	0.5100	0.2899	0.3600	
Efficiency	0.5136	0.589	0.0604	0.244	0.9397	0.5100	0.0588	0.2500	
Cum. Payoff	0.5136	0.589	0.0604	0.244	0.6501	0.5600	0.2899	0.3600	
Punish			0.3911	0.617			0.8498	0.4750	
p*-value\Ho:	NPR=PR		NPR<PR		NPPR=PPR		NPPR<PPR		
Contribution	0.0009		0.9630		0.0002		1.0000		
Efficiency	0.0193		0.8270		0.0696		0.7400		
Cum. Payoff	0.0193		0.8270		0.7055		0.5500		
p*-value\Ho:							PPA=PPR	PPA>PPR	
Contribution							first 6 rounds last 4 rounds	0.0039 0.0209	0.0000 0.0000

\*Wilcoxon-Mann-Whitney rank-sum test, two-sided and one-sided; compares mean group values NPA-no punishment with algorithm, NPR-no punishment random, PA-punishment with algorithm, PR-punishment random, NPPA-no punishment partial visibility algorithm, NPPR-no punishment partial visibility random, PPA-punishment partial vision algorithm, PPR-punishment partial vision

random.

In by period data efficiency in (1-1)>(2-1); (1-2)<(2-2); (3-1)=(4-1); (3-2)=(4-2). No difference in the efficiency between control groups and experimental groups in by group data.

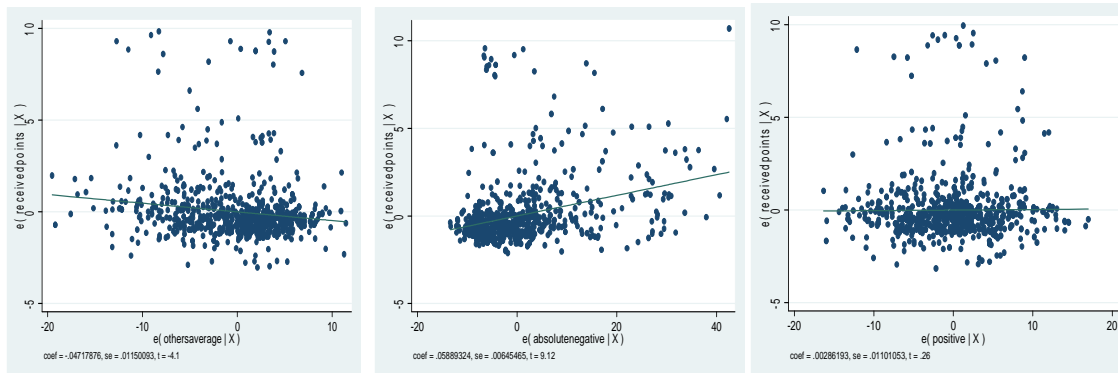
Table 6. Regression: determinants of getting punished

Independent variables	Dependent variable: received punishment points	
	Shuffle treatment	Partner treatment
Constant	1.9513 *** (0.49)	4.3965 *** (0.78)
Other's average contribution	-0.0472 *** (0.01)	-0.0922 *** (0.01)
Absolute negative deviation	0.0589 *** (0.01)	0.1236 *** (0.01)
Positive deviation	0.0029 (0.01)	-0.0511 *** (0.02)
	N=800	N=760
	F( 15, 784) = 9.76***	F( 21, 738)=37.26***
	Adj R-squared=0.14	Adj R-squared=0.50
	DW=2.17	DW=1.77

Note: Standard errors are in parenthesis. \*-significant at 10-percent, \*\*-significant at 5-percent, \*\*\*-significant at 1-percent or better level

Regression includes dummies for rounds and sessions in shuffle treatments and dummies for periods and groups in partner treatments

#### Shuffled treatments (design 3 and 4)



Others' average c

Absolute negative deviation

Positive deviation

Partner treatments (design1 and 2)

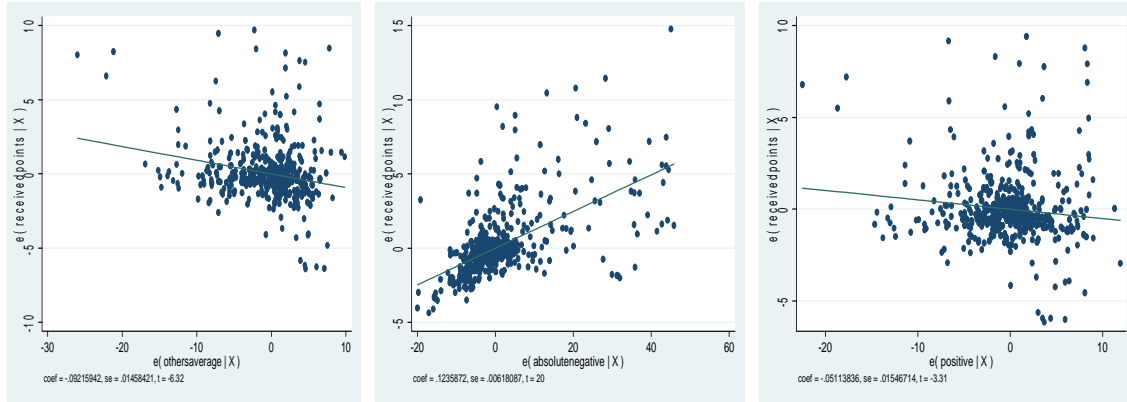


Table 7. Grid/Group scores in experimental vs. control groups

	grid		group		altgrid		altgroup	
	d1 vs. d2	d3 vs. d4	d1 vs. d2	d3 vs. d4	d1 vs. d2	d3 vs. d4	d1 vs. d2	d3 vs. d4
mean experimental	0.43	0.43	0.52	0.53	0.56	0.65	0.45	0.52
mean control	0.41	0.49	0.49	0.53	0.56	0.57	0.41	0.51
p*-value Ho: var1=var2	0.6734	0.0243	0.1632	0.8361	0.9707	0.1139	0.6632	0.9031
p**-value Ho: var1>var2	0.528	0.35	0.593	0.487	0.502	0.602	0.529	0.492

\*Wilcoxon-Mann-Whitney rank-sum test, two-sided

\*\*Wilcoxon-Mann-Whitney rank-sum test, one-sided

Table 8. Assurance game outcomes

	All	4AG	2AG
Mean payoff	2.79	2.38	3.20
st.dev	(2.12)	(2.22)	(1.93)
Mean efficiency	55.75	47.50	64.00
st.dev	(37.65)	(41.07)	(31.94)

Ho:

p-value* payoff Ho: 4AG=2AG	0.0425	
p-value* payoff Ho: 4AG<2AG	0.6110	
p-value* payoff Ho: S=R	0.2299	0.1250
p-value* payoff Ho: S>R	0.3930	0.4040
p-value* efficiency Ho: 4AG=2AG	0.0425	
p-value* efficiency Ho: 4AG<2AG	0.6110	
p-value* efficiency Ho: S=R	0.2299	0.1250
p-value* efficiency Ho: S>R	0.3930	0.4040

\* Wilcoxon rank-sum test, two-sided and one-sided  
S-sorted, R-random

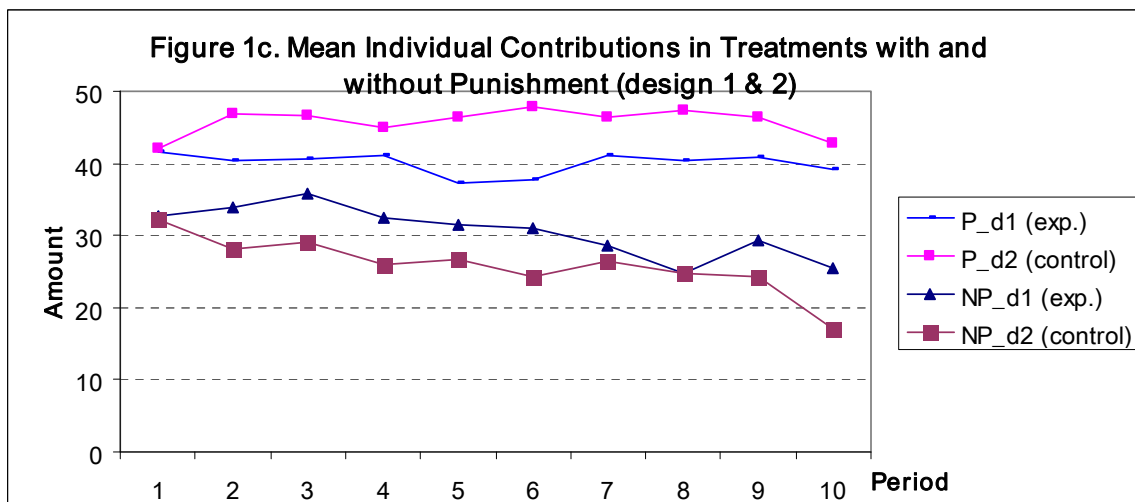
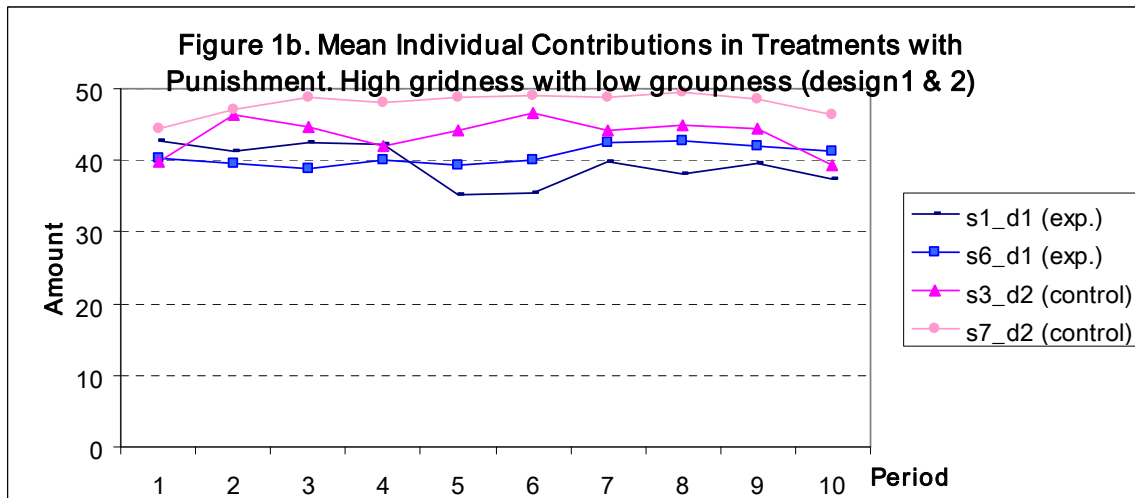
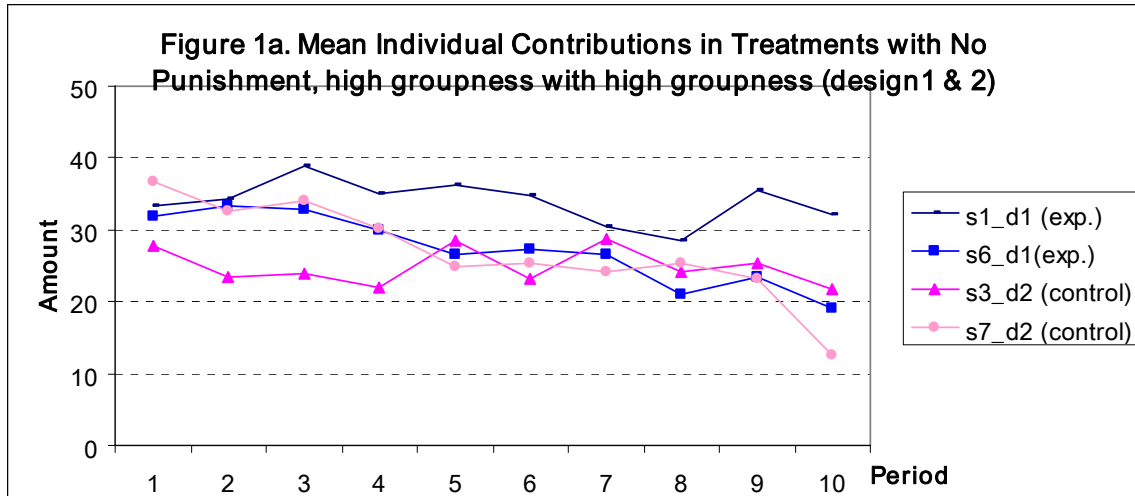
<b>Table9.</b>			
<b>Correlation between Choice and Gridness Score</b>			
	No Penalty	Penalty	Pooled
4 people group	0.093 (0.413)	0.270** (0.016)	0.169** (0.033)
2 people group	-0.047 (0.682)	0.176 (0.119)	0.062 (0.440)
Pooled	0.022 (0.784)	0.214** (0.007)	0.111** (0.048)
<b>Correlation between Choice and Groupness Score</b>			
	No Penalty	Penalty	Pooled
4 people group	-0.109 (0.335)	0.108 (0.924)	-0.054 (0.501)
2 people group	-0.219 (0.051)	-0.135 (0.233)	-0.174* (0.028)
Pooled	-0.163 (0.041)	-0.067 (0.397)	-0.115* (0.040)
Choice ``A``=1, Choice ``B``= 2; Result: High-groupness prefer choice ``A``; High-gridness prefer choice``B``			

**Table 5. Shuffle/Sorting /Pooled- Pearson Correlations**

		Shuffle=1		Sorting=1			All pooled			Shuffle=0, Sorting=1			Shuffle=1	
		(3-1) Contrib	(3-2) P Contrib	(3-2) Punish	(3-3) 4AG	(3-4) 2AG	Contrib	P Contrib	Punish	(1-1) Contrib	(1-2) P Contrib	(1-2) Punish		
Grid	Pearson's r	-0.086	-0.178**	-0.058	0.186	0.054	0.008	-0.146**	-0.026	0.011	-0.171**	-0.063		
	Sig. (2-tailed)	0.084	0.0003	0.249	0.103	0.637	0.745	0.000	0.300	0.823	0.001	0.206		
	N	400	400	400	80	80	1560	1560	1560	400	400	400		
Group	Pearson's r	0.072	0.073	-0.122*	0.076	-0.085	0.098**	0.007	-0.129**	0.050	-0.084	-0.259**		
	Sig. (2-tailed)	0.151	0.141	0.015	0.504	0.452	0.000	0.786	0.000	0.316	0.093	0.000		
	N	400	400	400	80	80	1560	1560	1560	400	400	400		
Alt Grid	Pearson's r	0.100*	0.044	-0.013	-0.047	-0.244*	0.076*	-0.074**	-0.034	-0.027	-0.267**	-0.045		
	Sig. (2-tailed)	0.045	0.383	0.794	0.676	0.029	0.003	0.006	0.175	0.594	0.000	0.368		
	N	400	400	400	80	80	1560	1560	1560	400	400	400		
Alt Group	Pearson's r	0.023	-0.084	0.048	0.0150	-0.120	0.026	-0.072**	0.044	-0.112*	-0.231**	0.023		
	Sig. (2-tailed)	0.649	0.092	0.341	0.896	0.289	0.305	0.004	0.083	0.026	0.000	0.654		
	N	400	400	400	80	80	1560	1560	1560	400	400	400		
		Shuffle=1		Sorting=0			All pooled			Shuffle=0, Sorting=0				
		(4-1) Contrib	(4-2) P Contrib	(4-2) Punish	(4-3) 4AG	(4-4) 2AG			AG	(2-1) Contrib	(2-2) P Contrib	(2-2) Punish		
Grid	Pearson's r	-0.079	-0.146**	0.061	0.150	0.071			0.111*	0.186**	0.011	-0.072		
	Sig. (2-tailed)	0.114	0.003	0.227	0.183	0.531			0.048	0.001	0.833	0.173		
	N	400	400	400	80	80			320	360	360	360		
Group	Pearson's r	0.057	0.132**	-0.020	-0.225*	-0.286*			-0.115*	0.285**	0.069	0.095		
	Sig. (2-tailed)	0.258	0.008	0.692	0.045	0.010			0.040	0.000	0.190	0.095		
	N	400	400	400	80	80			320	360	360	360		
Alt Grid	Pearson's r	0.092	0.258**	-0.142**	-0.171	-0.089			-0.128*	0.160**	-0.171**	0.173*		
	Sig. (2-tailed)	0.065	0.000	0.004	0.130	0.434			0.022	0.002	0.001	0.001		
	N	400	400	400	80	80			320	360	360	360		
Alt Group	Pearson's r	-0.082	0.079	0.124*	-0.181	-0.112			0.101	0.327**	0.195**	-0.064		
	Sig. (2-tailed)	0.103	0.113	0.013	0.108	0.324			0.070	0.000	0.001	0.225		
	N	400	400	400	80	80			320	360	360	360		

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).



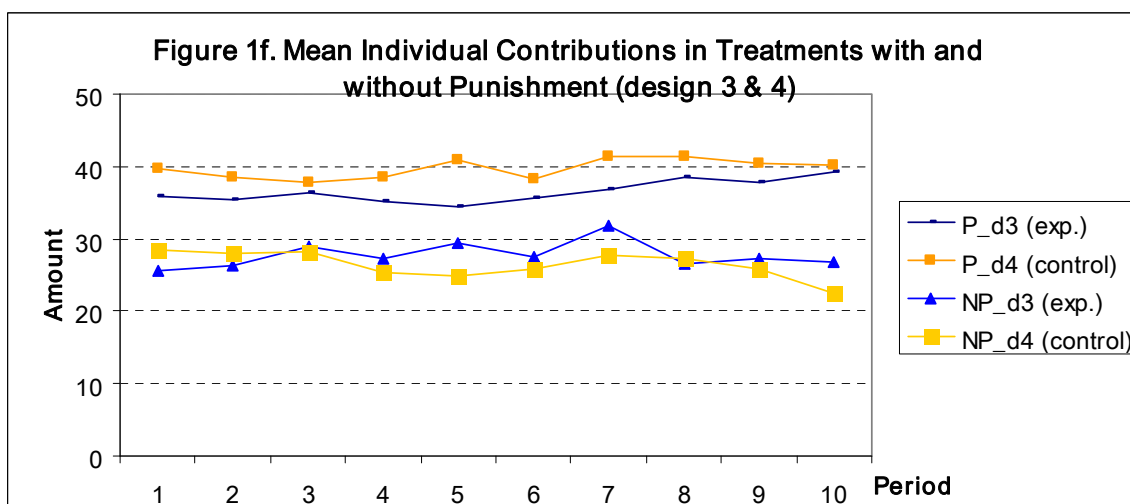
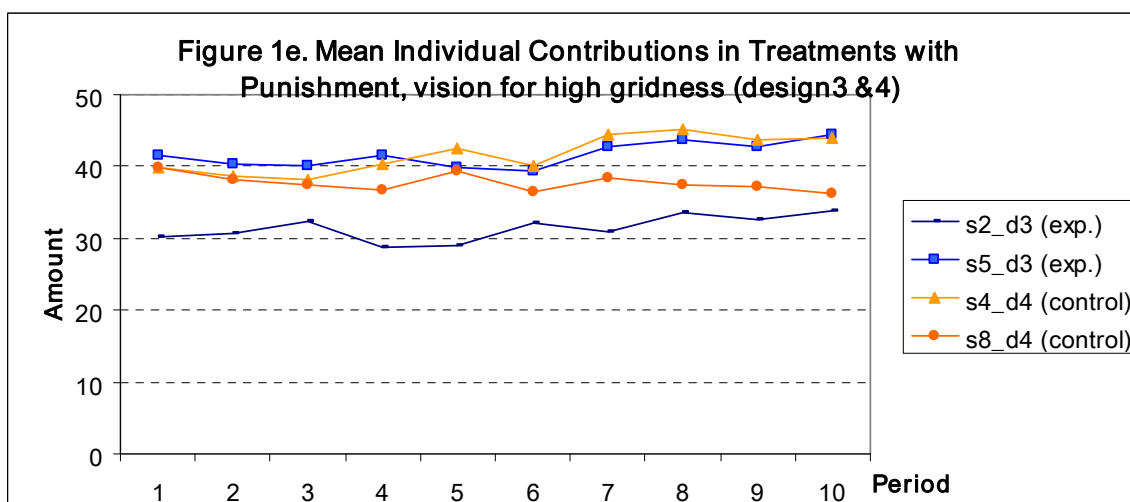
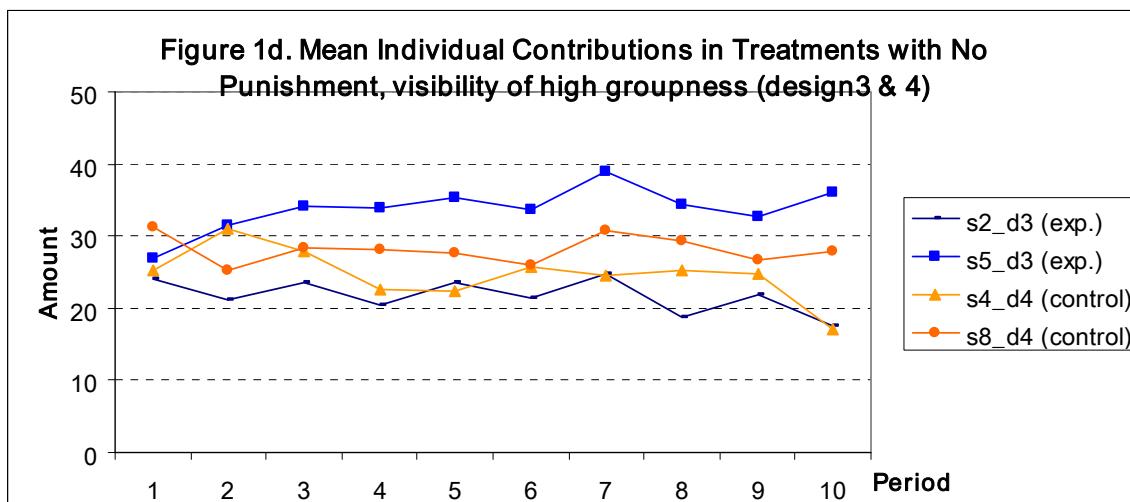


Figure 2. Mean Contributions by Typology  
in No Punishment and Punishment treatments

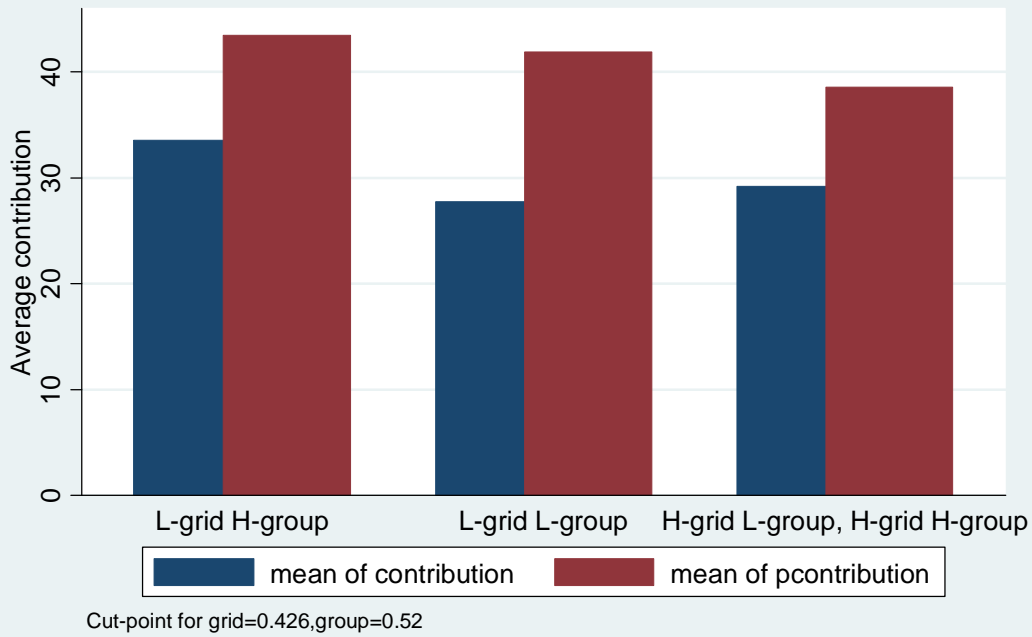
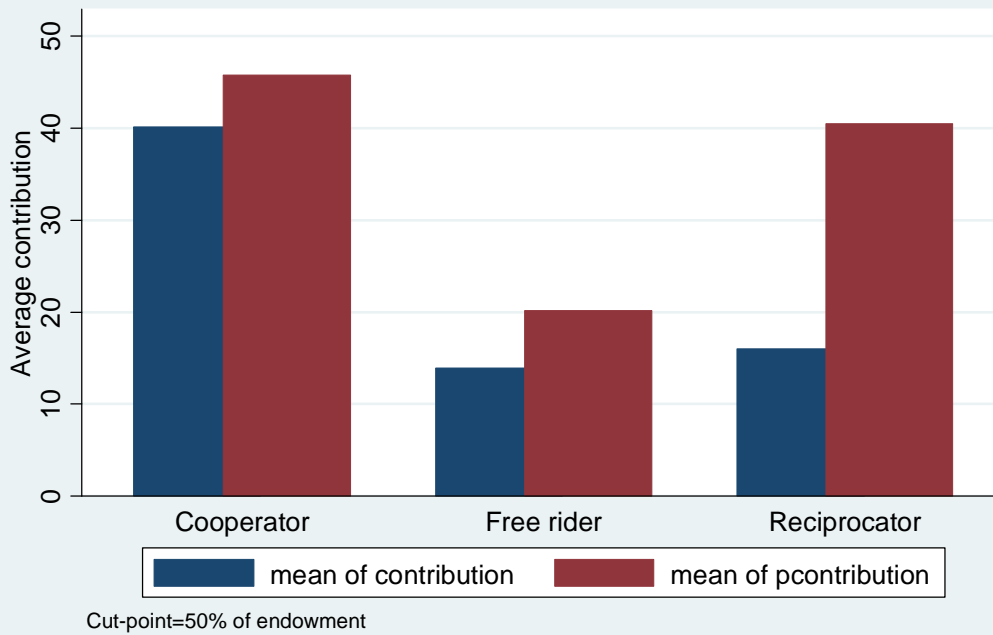


Figure 3. Mean Contributions by Type  
in No Punishment and Punishment treatments



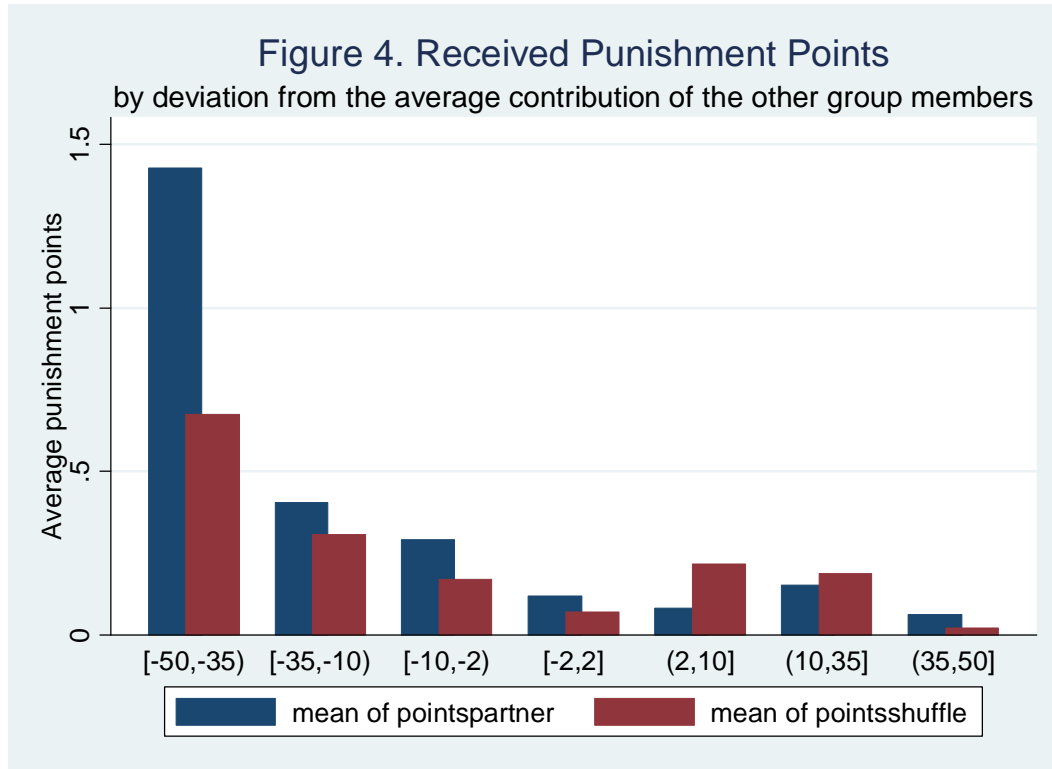


Figure 5. Coordinative Compensation Games:

4-person Assurance Game 1:

	A	B
A	5,5	0,3
B	3,0	3,3

4-person Assurance Game 2:

	A	B
A	5, 5	0, 3 - #(A)
B	3 - #(A), 0	3 - #(A), 3 - #(A)

2-person Assurance Game 3:

	A	B
A	5,5	0,4
B	4,0	4,4

2-person Assurance Game 4:

	A	B
A	5,5	0,2
B	2,0	4,4