Conventions for Selecting Among Conventions An Evolutionary and Experimental Analysis

Susanne Büchner,[*] Werner Güth[†]and Luis M. Miller[‡]

May 2007

Abstract

Conventions can be narrowly interpreted as coordinated ways of equilibrium play, i.e., a specific convention tells all players in a game with multiple strict equilibria which equilibrium to play. In our view, coordination often takes place before learning about the games. Thus, one has to coordinate on a prescribing principle of equilibrium selection. For the subclass of 2x2-bimatrix games with two strict equilibria we analyze the evolutionary stability of various such principles. In our experiment, we allow participants to first coordinate before playing various games. Based on between-subjects treatments, participants do this behind a complete (they know neither their role nor the game parameters), a partial (they know either their role or the game parameters) veil of ignorance, or with no ignorance (they know their role and the game parameters).

Keywords: coordination games, conventions, experimental economics, evolutionary stability.
JEL: C72, C91

[*]Max Planck Institute of Economics, Strategic Interaction Group, Kahlaische Str. 10, 07745 Jena, Germany; email: buechner@mpiew-jena.mpg.de.

[†]Max Planck Institute of Economics, Strategic Interaction Group, Kahlaische Str. 10, 07745 Jena, Germany; email: gueth@mpiew-jena.mpg.de.

[‡]Max Planck Institute of Economics, Strategic Interaction Group, Kahlaische Str. 10, 07745 Jena, Germany; email: miller@mpiew-jena.mpg.de and Consejo Superior de Investigaciones Científicas, Instituto de Estudios Sociales Avanzados de Andalucía, Campo santo de los mártires 10, 14004, Córdoba, Spain; email: lmmiller@iesaa.csic.es.

1   Introduction

If we had to coordinate anew with others how to behave in each of our many possible encounters, we would be overburdened by coordination effort. To save such effort, we rely on general principles of coordination, like priority when waiting for a resource, such as a medical treatment; when driving on highways; when working together, e.g., when several police units are needed to solve a problem.[1]

There are not only multiple actions, like using the left- or right-hand side of the road, but also multiple principles of coordination, like giving priority to the first or last arriving guest when we want to be served. We thus have to distinguish between:

- *specific conventions*, which tell us which action to choose, e.g., when playing a game, and

- *general conventions*, which allow us to derive specific conventions for each specific instance of interaction, e.g., a game to be played.

Characteristics of a specific convention are arbitrariness and stability. Conventions are arbitrary, because there is at least one alternative pattern of social behavior, and stable due to their self-enforcing character.

In the last few decades, several authors have tried to explain two important aspects of specific conventions. They have explored the conditions under which a specific convention could emerge (Lewis 1969; Schotter 1981; Sugden 1986). A growing literature analyzes the evolution of specific conventions (Young 1993; 1996; 1998). However, the problem of how to coordinate on a general convention when there is a multiplicity of conventions seems bew, to the best of our knowledge.

Our study is an attempt to study coordination on a general convention for selection between specific conventions. Such general principles will be understood as cognitive devices to coordinate expectations (Nozick, 1993: 5). General conventions make actions predictable and, therefore, by using them, individuals have a better chance of coordinating their actions. Ullmann-Margalit (1977) offered three arguments regarding the usefulness of this type of general coordination principles:

- first, "a [coordination] norm is capable of regulating and channeling the expectations [...] of anonymous participants" (p. 85);

---

[1] For instance, in Spain when units from different police forces have to act in cases of terrorism or organized crime they need to work coordinately, not only as a question of effectiveness but also of efficiency, on the allocation of resources and efforts. Therefore, the first preference of both groups is to act together. But in order to act together, they need to agree on the principle that is going to lead future interactions, and they may have different preferences regarding the coordination principle. In this particular case, the only principle that is legally enforced (as one of the possible answers to solve recurrent coordination problems among different police forces) is priority when arriving to the particular case (Miller 2007) .

- second, "while a regularity extracted from past events might sometimes be continued in more than one way, a norm will provide the principle of continuation which will resolve potential ambiguities in most future events" (p. 87); and

- third, "there is a higher degree of articulation and explicitness associated with a norm than with a mere regularity of behavior" (p. 87).

Here we will explore how general conventions[2] can emerge when prescribing specific conventions in strategic situations. More specifically, we study both evolutionarily and experimentally how establishing general conventions enhances the emergence of specific conventions and, therefore, the level of coordination.

General conventions will be analyzed evolutionarily by assuming that individuals are committed to the respective behavior when deciding in a specific game. In our experiment, participants can try to establish conventions where between-subjects treatments differ in what subjects know about their individual choice problems when suggesting which principle should drive behavior. Section 2 introduces the basic game class and the general principles we are considering. Section 3 outlines an evolutionary analysis of general conventions. Section 4 describes the experimental design and procedure. Section 5 presents the experimental results. Section 6 concludes.

2 Game class and general principles

The class of games we are considering is a subclass of 2x2-bimatrix games (two players, each facing two different actions). Four payoff parameters a, b for player B and c, d for player A, satisfying $1 > a > b > 0$ and $1 > c > 0 > d > -1$, determine the payoff matrix:

$$
\begin{array}{c|c|c|}
 & X_B & Y_B \\
\hline
X_A & 1,a & d,0 \\
\hline
Y_A & 0,b & c,1 \\
\hline
\end{array}
$$

with payoffs in the usual order. In addition to the two strict equilibria $X = (X_A, X_B)$ and $Y = (Y_A, Y_B)$, there exists a mixed strategy equilibrium (p^*, q^*) with $p^* = \Pr\{X_A\}$, $q^* = \Pr\{X_B\}$ and payoffs $U_i(p^*, q^*)$:

$$
p^* = \frac{1-b}{1-b+a} \quad , \quad q^* = \frac{c-d}{1+c-d}
$$
$$
U_A(p*, q*) = \frac{c}{1+c-d} \quad , \quad U_B(p*, q*) = \frac{a}{1-b+a} \quad .
$$

It is natural (Harsanyi and Selten 1988) to start by considering first the two strict equilibria as solution candidates. Due to the coexistence of X and Y as strict equilibria, this alone would not resolve strategic uncertainty. Thus, even this most refined equilibrium idea (van Damme 1987) does not help players to

---

[2]In the following, we are going to use the terms general convention and general principle with the same meaning.

solve their coordination problem. This becomes even worse when the attractiveness of strict equilibria is not commonly known and accepted. Other principles as well as other candidates might then become relevant.

A selection principle would solve the coordination on how to play one of the games described above, e.g., in the sense of recommending a (strict) equilibrium or, more generally, one of the four possible pure strategy vectors. Such a coordination requires, of course, that both players are commonly aware which coordination principle should be applied. In the experiment, we will allow the two interacting participants to communicate, at least to some degree, how they evaluate certain principles of coordinating on strategy selection, although this does not guarantee commonly known true rankings of such principles.

The most convincing concept seems to be *Risk dominance* (R) which, for the class of games considered, is characterized by the three axioms of *best reply invariance*,[3] *isomorphic invariance*,[4] and *monotonicity*[5] (Harsanyi and Selten 1988). By *best reply invariance* the game can be transformed into:

	X_B	Y_B
X_A	$1, a$	$0, 0$
Y_A	$0, 0$	$c - d, 1 - b$

This game can, in turn, be transformed by *isomorphic invariance*, actually by affine utility transformation, into:

	X_B	Y_B
X_A	$1, \frac{a}{1-b}$	$0, 0$
Y_A	$0, 0$	$c - d, 1$

Thus, due to *monotonicity*, the solution is:

- $X = (X_A, X_B)$ if $a > (c - d)(1 - b)$ and

- $Y = (Y_A, Y_B)$ if $a < (c - d)(1 - b)$

where we will guarantee, by our parameter choices in the experiment, applicability in the sense of

$$a \neq (c - d)(1 - b).$$

Next, we will consider *Loss avoidance* (L), which excludes player A's choice of strategy X_A since it, together with Y_B, yields a loss for player A. *Loss avoidance* plus equilibrium behavior thus suggests the solution (Y_A, Y_B) always.[6]

---

[3] Since, according to rationality, one should always best respond according to what one expects others to do, one only preserves the information to which strategy constellations strategies are best responding. Note that best replay preserving transformations do not preserve payoff dominance.

[4] Allowing to change names of players, strategies, and affine utility transformations.

[5] Increasing one player's payoff for an equilibrium candidate will render this candidate dominant if it has been undominated before.

[6] For previous experimental results regarding *Loss avoidance* as an equilibrium selection principle, see Cachon and Camerer (1996) and Rydval and Ortmann (2005).

Unbiasedness (U), on the other hand, assumes that the respective other player uses both strategies with the same probability and that one reacts optimally to this expectation. Consequently:

- player A chooses $\begin{cases} X_A & \text{if} \quad 1+d > c \\ Y_A & \text{if} \quad 1+d < c \end{cases}$

- player B chooses $\begin{cases} X_B & \text{if} \quad a+b > 1 \\ Y_B & \text{if} \quad a+b < 1 \end{cases}$

Again our parameter choices guarantee $1+d \neq c$ and $a+b \neq 1$.

Minimizing risk (M) suggests the strategy for which the worst result is better. Thus, player A should select Y_A and player B the strategy X_B, what altogether does not yield self-enforcing expectations.

Sum maximization (S) simply recommends the strategy constellation yielding the largest payoff sum, i.e.,

- (X_A, X_B) if $a > c$

- (Y_A, Y_B) if $a < c$

where we again exclude the degenerate case $a \neq c$.

Generosity (G) ask each player to choose what is better for the other when he or she uses both strategies with equal probability as for (U). Thus, player A should use Y_A and player B the strategy X_B as according to criterion (M). This illustrates that the same selection criteria may be justified in several ways.

Indifference induction (I) suggests to randomize in such a way that the other player does not care what to choose. This, of course, suggests the mixed strategy–equilibrium (p^*, q^*). Note that the two transformations, used for deriving the risk dominant equilibrium, did not change (p^*, q^*), revealing the best reply invariance of the two transformations.

Finally, *Alternating* (A) recommends to alternate both strategies, e.g., choosing X_i in the even rounds and Y_i in the odd ones.

One may argue that there exist other principles for selecting choices, e.g., do the worst for your co-player. We do not want to include all of them but only those with some normative appeal. Given this limitation, our list of "normative principles" seems quite comprehensive.

3 Evolutionary analysis

The aim of this section is to study the evolutionary stability of several general conventions when playing the basic game presented above. Although we rely on a special class of games,[7] our results are general in the sense that we do not perform for numerically specified but rather the general type of such games.

---

[7] An evolutionary analysis of a habitat allowing for different types of games is provided by Güth and Napel (2006).

Here we are not concerned about the evolution of specific conventions,[8] i.e., of strategies describing how to play in a coordination game, but about the evolution of general principles that recommend specific conventions. Whereas specific conventions are ways to play coordination games, general principles tell us which specific convention to use in certain classes of coordination games.

To do this, let us assume that individuals have types corresponding to the above elaborated general conventions R, U, L, M, S, G, I, and A. An M-type or M-individual will thus always choose Y_A as player A and X_B as player B, whereas for a U-individual behavior will depend on the game parameters: as player A the U-individual will use X_A (resp. Y_A) if $1 + d > c$ (resp. $1 + d < c$) and as player B will use X_B (resp. Y_B) if $a + b > 1$ (resp. $a + b < 1$). Assuming that each individual plays the game in both roles (one-population-model) we can analyze which monomorphic populations (all individuals are of the same type) is evolutionarily stable.

The answer will, in general, depend on the game parameters. Let us say that a T-monomorphism with $T \in \{R, U, L, M, S, G, I, A\}$ is an evolutionarily stable general convention if there exists a generic parameter region where T is the unique best reply to itself. This would immediately eliminate M and G as stable monomorphisms since they imply the same specific convention. To avoid such a trivial result, we therefore merge M and G as M/G and refer to M/G as a unique general principle.

To analyze evolutionary stability we have to derive for each type $T \in \{R, U, L, M/G, S, I, A\}$, the success of a type $\widetilde{T} \in \{R, U, L, M/G, S, I, A\}$ entering as a $\widetilde{T}$-mutant a T-monomorphic population. For $T = U$ and $\widetilde{T} = U$ the success is, for instance,

- $1 + a$, if $1 + d > c$ & $a + b > 1$ since then U-types play X_i in both roles i=A, B ,

- d if $1 + d > c$ & $a + b < 1$ since then U-types as player A use X_A but Y_B as player B,

- b if $1 + d < c$ & $a + b > 1$ since then U-types as player A use Y_A but X_B as player B,

- $c + 1$ if $1 + d < c$ & $a + b < 1$ since then U-types play Y_i in both roles i=A, B.

Assume now for $T = U$ a $\widetilde{T} = M/G$-mutant. Proceeding as above, the success of the $\widetilde{T} = M/G$-mutant is

- a if $1 + d > c$ & $a + b > 1$ since the U-type plays X_i always and the M/G-type Y_A (resp. X_B),

- $c + a$ if $1 + d > c$ & $a + b < 1$ since the U-type uses X_A resp. Y_B and the M/G-type Y_A (resp. X_B),

---

[8] A complete model of the evolution of specific conventions was presented by Young (1993).

- b if $1 + d < c$ & $a + b > 1$ since the U-type uses Y_A resp. X_B and the M/G-type Y_A (resp. X_B),

- $c + b$ if $1 + d < c$ & $a + b < 1$ since the U-type plays Y_i always and the M/G-type Y_A (resp. X_B).

Proceeding in this way for all constellations $T, \widetilde{T} \in \{R, U, L, M/GS, I, A\}$, we can derive the complete success table (Appendix A).

For U to qualify as a general convention, we would therefore have to find a generic parameter region where $R(U, U) > R(T, U)$ for all $T \neq U$ and $T \in \{R, U, L, M/G, S, I, A\}$. In case of $a + d > c$ & $a + b > 1$ one has, for instance, $R(U, U) = 1 + a > a = R(M/G, U)$, meaning that a U-monomorphic population cannot be invaded successfully by M/G-mutants if only games in the generic subregion $1 + d > c$ & $a + b > 1$ are played.

Proceeding in this way for the whole table, we obtain the following results:

Result 1: We can rule out I^* and M/G^* as general conventions. These selection principles are not evolutionarily stable strategies (ESS) because they are no best replies to themselves in any region of the table.

Result 2: There are several general conventions which are ESS (see table 1).

Result 3: There is generic multiplicity of general conventions. For $(c - d)(1 - d)(1 - b) < a < c$ & [not: $1 + d > c$ & $a + b > 1$], for instance, R^* and L^* are evolutionarily stable strategies (ESS). In addition, A^* is an ESS for all parameter regions.

Table 1: **Evolutionarily stable general conventions**

ESS	Parameter subregions
U*	$1 + d > c$ & $a + b > 1$ & $a < c$ & $a < (c - d)(1 - b)$
S*	$(c - d)(1 - b) > a > c$
R*	$(c - d)(1 - d)(1 - b) < a < c$ & not: $[1 + d > c$ & $a + b > 1]$
L*	$c > a > (c - d)(1 - b)$
A*	For all parameter regions

The generic multiplicity of evolutionarily stable general conventions reveals that analyzing the evolution of general rather than specific conventions can shift the coordination problem only to the deeper level of coordinating principles rather than behavior. So the major step forward is mainly that the evolutionary justification of general conventions assumes that we have evolved how to play by learning principles and not by learning game playing. Our results, furthermore, indicate that not all principles can be justified in that way and may therefore not be used in the experiment.

4 Experimental design and procedure

For the experiment, 160 undergraduate students were recruited (from different disciplines) at Jena University, using ORSEE 2.0 (Greiner 2004). Five exper-

imental sessions (every session corresponding to a different treatment) were conducted, each using a different group of 32 participants. Instructions of the experiment can be found in Appendix B. The experiment was computerized using z-Tree (Fischbacher 1999). Subjects received written instructions, which were also read aloud by the experimenter to ensure everyone understood them. No communication other than through the use of principles (as we will explain below) was permitted. Subjects could not identify which members of their matching group they were actually interacting with. At the end of the experimental session, subjects were paid in cash according to their payoff in the game (plus a show-up fee of 2 €).

Every subject was assigned a role (A or B), and everyone interacted in a group of four subjects playing 20 single games. There were ten different games,[9] which played twice. During the first ten games (1-10), they played with one partner and during the second ten (11-20) with another partner. Every treatment differed in the manipulation of three variables: (i) knowledge of the actual game parameters, (ii) knowledge of one's own role, and (iii) the strategic use of principles. Our evolutionary analysis is, of course, restricted to treatments N and NC. A similar analysis for treatments C, P, and PC is possible but would require quite a bombastic Bayesian setup (expectations concerning role assigment as well as parameters specification). Here we distinguish these five treatments to test whether general principles are used game specifically or rather as general routines allowing coordination without knowing the details of the game. Table 2 reports the main characteristics of the experiment.

Table 2: Characteristics of the treatments

	Tr-C	Tr-P	Tr-PC	Tr-N	Tr-NC
Number of sessions	1	1	1	1	1
Periods per session	20	20	20	20	20
Matching groups	8	8	8	8	8
Participants	32	32	32	32	32
Kparameters	No	No	Yes	Yes	Yes
Kroles	No	Yes	No	Yes	Yes
Use of principles	Yes	Yes	Yes	Yes	No
Coordination	69.38%	53.75%	78.13%	74.38%	66.56%
Average earnings	7.5€	6.8€	8.5€	8.2€	7.6€

In every treatment, participants had to fulfill two tasks. First, they had to rank the principles and, second, they had to play ten successive games under different information conditions, captured by between-subjects treatments. The first task implied to fill in a table with the proposed principles (see the ranking screen in Appendix C). They did not have to suggest a complete order; the only requirement was to select at least one principle. After every subject had given their ranking, they were matched with their partner with whom they played the ten rounds. This was repeated with a new partner. Before playing the games,

---

[9] The ten games can be consulted in the instructions (see Appendix B).

they knew both their own ranking and the ranking of their partner (see the decision screen in Appendix C). This condition was only changed in the control group. No information regarding the outcomes was provided after each round. At the end of the twenty periods they were informed about their earnings and were paid in cash (see the average earnings per subject and treatment in table 2).

Every session followed the same protocol, and treatments only differed in the three variables that we manipulated (see table 2). In treatment C, participants made decisions under a *complete veil of ignorance* because they knew neither the actual game parameters nor their own role. In treatments P and PC, they interacted under a *partial veil of ignorance* since in treatment P they knew their own role but not the actual game parameters, while in treatment PC they knew the actual game parameters but did not know their own role. Treatment N supposed *no veil of ignorance* at all since participants knew game parameters and roles. Treatment NC was a control group in which they had to rank, but did not communicate rankings to their partners.

5   Experimental results

In the following, we mainly try to answer four questions:

First, we analyze which principles were most frequently used. To this end, we present: first, data regarding the first-ranked principles and then data regarding all rankings.

Second, we test consistency with first-ranked principles, i.e., whether actual behavior is consistent with these.

Third, we investigate whether using general principles improves coordination by comparing our results with those from previous experiments.

Fourth, we explain the differences between treatments by the two treatment variables that we manipulated, namely: knowing one's role and/or the actual game parameters.

5.1   Which principles were chosen?

Figure 1a plots the frequency of first-ranked principles. Principles that are not evolutionarily stable strategies (ESS) were rarely selected: no subject ranked *Minimizing risk* first, one subject *Generosity*, and ten subjects *Indifference Induction*. Altogether, non-evolutionarily stable principles were ranked first in less than 7% of the cases.

The other principles were ranked first by a larger number of subjects. The only exception seems to be principle U (*Unbiasedness*) which was selected only twice as first principle. One possible explanation is that coordinating on general convention U does not generate a specific convention (equilibrium) since U assumes reacting optimally to the other player using both strategies with equal probability. In other words, principle U did not provide "a system of suitably concordant mutual expectations" (Lewis 1969: 25).

Among the other four general conventions, the most frequent principle was *Sum maximization* (*S*), followed by *Loss avoidance* (*L*), *Alternating* (*A*), and *Risk dominance* (*R*). *S* is suggested by concerns for efficiency, *L* by frustation aversion, *A* by procedural fairness, and *R* by comparing equilibrium candidates against each other. All these criteria pointed out a particular specific convention. There are no statistical differences among treatments regarding which principle was ranked first (Kruskal-Wallis test; $X^2 = 1.735, d.f. = 4, p = 0.784$). What explains the popularity of *Sum maximization*? Besides its obvious simplicity and easy applicability, it offers a welfare measure which neglects the distribution conflict (who gets more). This might have been less negligible if participants were playing only one round or were being paid only for a randomly selected round.
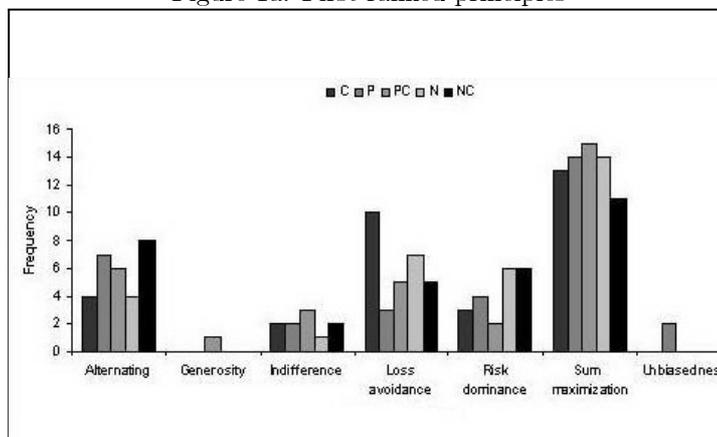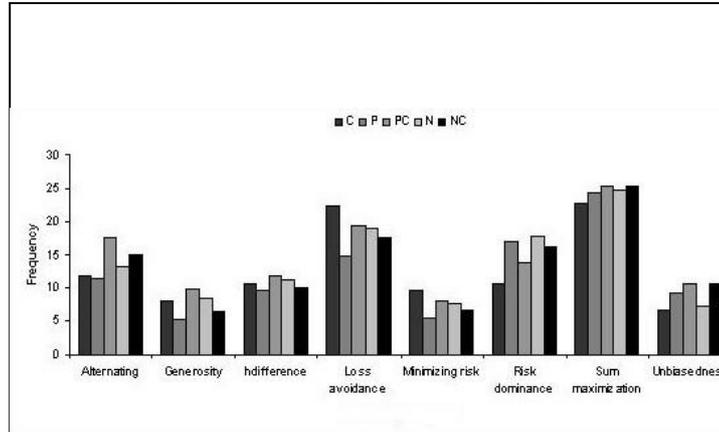
Figure 1a: First-ranked principles

Figure 1b, instead of considering only first-ranked principles, relies on the all rankings. When weighting principles by the inverse of their ranks, the results are very similar to the previous figure: principles that did not provide "mutual and concordant expectations" are used less, and the order of the remaining ones was as follows: *S*, *L*, *R*, *A*. Again, there are no statistical differences among treatments regarding the form in which principles were ranked (Kruskal-Wallis test: $X^2 = 4.858$, d.f. $= 4$, $p = 0.784$, for principle A; $X^2 = 5.722$, d.f. $= 4$, $p = 0.221$, for principle G; $X^2 = 2.074$, d.f. $= 4$, $p = 0.722$, for principle I; $X^2 = 4.759$, d.f. $= 4$, $p = 0.313$, for principle L; $X^2 = 5.367$, d.f. $= 4$, $p = 0.252$, for principle M; $X^2 = 4.614$, d.f. $= 4$, $p = 0.329$, for principle R; $X^2 = 1.258$, d.f. $= 4$, $p = 0.869$, for principle S; $X^2 = 5.228$, d.f. $= 4$, $p = 0.265$, for principle U).

Figure 1b: All used principles weighted

5.2 Is behavior consistent with first principles?

Figure 2 represents the percentage of consistency with the first-ranked principle, that is, whether participants relied on the strategy prescribed by their first ranked-principle. We have only considered cases with unambiguous first-ranked principles. So the data refers to 113 participants only. We also neglect principles that are not ESS and principle U that was nearly never chosen. The level of consistency with principles ranges from 66.79% for principle R (*Risk dominance*) to 85.5% for principle S (*Sum maximization*). The average of consistency with principles is, at 78.19%, quite high. There are no statistical differences among treatments regarding the level of consistency with principles (Kruskal-Wallis test; $X^2 = 4.144$, d.f. $= 4$, $p = 0.387$).

Figure 2: Consistency with first-ranked principles
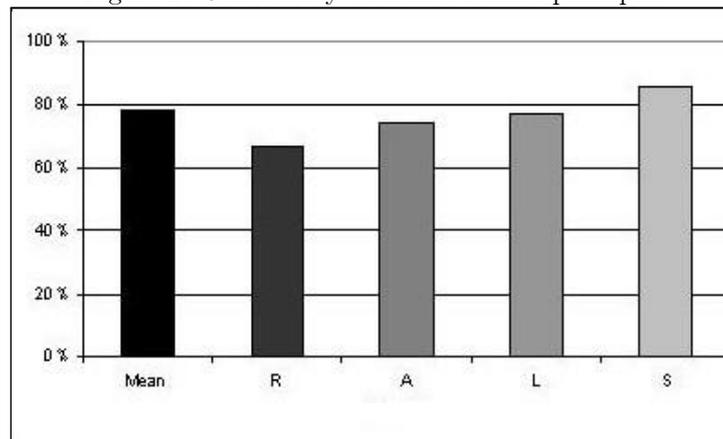
In our view, these data show that subjects use general principles as *cognitive devices* to achieve high coordination levels and to avoid reasoning about their

11

strategy for every single instance of interaction. This is related to the so-called *cognitive answer* to the explanation of norms (Sacconi and Moretti, 2002). But not all the principles play the same role in our experiment. In particular, the level of consistency with principle S (*Sum maximization*) is statistically different from the other principles for every binary comparison between principle S and the other three principles (Mann-Whitney U-test: $z = -2.496$, $p = 0.013$, 2-tailed, for comparison between principles A and S; $z = -2.414$, $p = 0.016$, 2-tailed, for comparison between principles L and S; $z = -3.888$, $p < 0.001$, 2-tailed, for comparison between principles R and S). This result is surprising since S is an efficiency criterion, and several experimental studies found no efficiency concerns in coordination games (Van Huyck, Battalio, and Beil 1990; 1991; Van Huyck, Cook, and Battalio 1997 who, however, rely on symmetric games). According to our results, efficiency might be an attainable end when subjects can communicate *ex ante* that they are going to use such a principle.

## 5.3 Does using principles improve coordination?

Several experimental studies have shown that the coordination achieved in standard versions of the battle of the sexes game without communication is between 40% and 50% (see Cooper *et al.* 1989; 1993; Straub 1995). Different devices have been tried to increase this level of coordination, like offering an outside option to one of the subjects, permitting different levels of communication between the players, or introducing a particular timing of the decision-making process (see Camerer 2003 for a review). However, to the best of our knowledge, no experiment has been run so far in which participants could choose a particular general principle before playing the coordination games. The mean coordination level in our experiment is 68.78% and increases from 53.75% in treatment P to 78.13% in treatment PC (see figure 3). There are significance differences between treatments (Kruskal-Wallis test; $X^2 = 12.667$, d.f. $= 4$, $p = 0.013$). Comparing our results with those obtained by other researchers,[10] we conclude that using principles improves coordination levels.

We run our control treatment to test this hypothesis. As explained before, in the control treatment (NC) subjects had to rank principles but did not communicate their ranking to their partner. If we compare treatments N and NC, which differ only in the strategic use of principles, we find that the coordination level achieved is 8% lower in treatment NC (Mann-Whitney test; $z = -2.088$, $p = 0.038$, 2-tailed). Thus, being able to communicate one's convention of choosing (specific) conventions definitely helps.

---

[10] In an experimental 2-players battle of the sexes game in which participants could communicate (cheap talk) the strategy that they were going to choose, Cooper *et al.* (1989) obtained a coordination level of 55%. In a replication of this treatment (Cooper *et al.*, 1993), the coordination level achieved was 58%.

Figure 3: Coordination by treatment
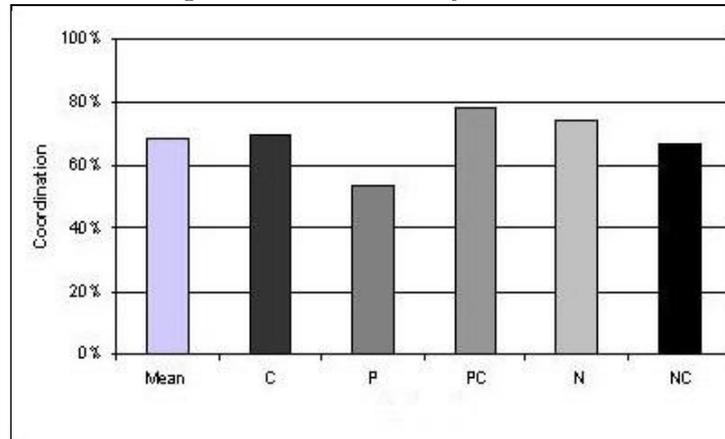
Did subjects maintain the same coordination patterns over 20 rounds of play? Figure 4 plots the aggregate coordination levels achieved for groups of five rounds each (periods 1-5, periods 6-10, periods 11-15, and periods 16-20). What we observe are mainly two aspects: First, coordination in the second part of the experiment (periods 11-20) is significantly higher than coordination in the first part (periods 1-10) (Wilcoxon signed ranks test; $z = -2.858$, $p = 0.004$, 2-tailed).[11] Second, in every treatment (except in C), coordination is lower in the third phase than in the second one, due to the change of partner and the necessity to coordinate on principles again, i.e., to a so-called restart effect. In general, we can conclude that coordination improves with experience.

Figure 4: Coordination by group of periods
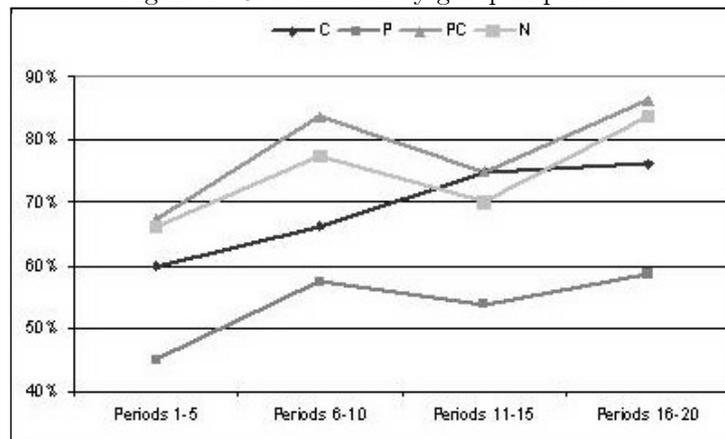
[11] In this and the next section, we neglect data from treatment NC (the control group) in order to control for the effect of the strategic use of principles.

5.4 How to explain treatment effects?

When analyzing the differences between treatments, we find that the two treatment variables that we manipulated affect coordination differently. Whereas knowing the actual game parameters improves coordination, knowing one's own role lowers coordination levels. This clearly demonstrates the need for studying coordination in assymmetric games rather than focusing on symmetric ones as in previous experimental studies.

When participants know the actual game they are playing, they can better predict the others' behavior and, in fact, this is what has happened. As we can see in table 3, the mean coordination among those who knew the numerical parameters of the game is almost 15% higher than the mean of those who did not (Mann-Whitney test; $z = -2.897$, $p = 0.004$, 2-tailed).

Table 3: Coordination by information

		Kroles	
		YES	NO
Kparameters	YES	74.38%	78.13%
	NO	53.75%	69.38%

Not knowing the roles is different. Although there is a tendency to achieve better coordination results in those treatments where the roles are unknown, the behavior of those who knew their role is not statistically different from the behavior of those who did not (Mann-Whitney test; $z = -1.212$, $p = 0.226$, 2-tailed). Especially, when one knows who has which role but not yet the game parameters, coordination seems to be most difficult, even when participants can communicate conventions for choosing specific conventions.

6 Conclusion

Interpreting a specific convention as coordinated strict equilibrium play in coordination games, we were interested in understanding whether and, if so, how communicating principles of equilibrium selection can improve coordination in a multiplicity of coordination games. When ranking such principles, participants were more or less informed about their specific situation according to a 2x2 factorial design. They did (not) know their role and/or the game parameters of the asymmetric coordination game. Our findings, which are rather encouraging, are as follows:

- Regardless of the veil of ignorance, participants are primarily suggesting an efficiency principle.

- Most participants feel obliged by their first-ranked principles.

- Communicating principles improves coordination.

- Knowing one's role in an asymmetric coordination game and not knowing its payoff parameters is detrimental to coordination.

For the institutional design, these findings can already be very helpful, for example when to decide on organizational coordination principles. In our view, there is no doubt that we rely on general principles to avoid coordinating anew in each interaction. But given that we do so, we better learn how this is done and how much it helps to coordinate behavior.

More specifically, the strong support of the *Sum maximization principle* as, for instance, illustrated in figure 1 seems to be quite comforting from a welfare-theoretic perspective. Rather than focusing on who gets what, people seem to be guided by the consideration "do what in sum or on average is best." Unfortunately, this may not extend to rare decisions with large stakes where it matters crucially which equilibrium is played. In such rare situations, one may tend to select between strict equilibria rather than, engaging in preplay, bargaining them by general principles which disregard the specific situation at hand.

But often coordination games are low-cost decision-making situations, e.g., who pays a minor bill, who provides a minor service, etc. (Kliemt 1986; Kirchgässner and Pommerehne 1993). Relying on welfare-enhancing general principles in such cases may be important, not because of what happens in a very specific encounter but simply because we face a lot of such low-cost/stake coordination games for which we rely on general conventions for selecting specific ones.

References

[1] Cachon, G. P. and C. F. Camerer, 1996. Loss avoidance and forward induction in experimental coordination games. Quarterly Journal of Economics 111, 165-194.

[2] Camerer, C. F., 2003. Behavioral Game Theory. Princeton, Princeton University Press.

[3] Cooper, R., D. V. DeJong, R. Forsythe and T. W. Ross, 1989. Communication in the Battle of the Sexes Game: Some Experimental Results. RAND Journal of Economics 20 (4), 568-587.

[4] Cooper, R., D. V. DeJong, R. Forsythe and T. W. Ross, 1993. Forward induction in the Battle of the Sexes Game. American Economic Review 83 (5), 1303-1316.

[5] Fischbacher, U., 1999. z-Tree-Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, Working Paper nº 21, Institute for Empirical Research in Economics, University of Zurich.

[6] Güth, W. and S. Napel, 2006. Inequality Aversion in a Variety of Games. An Indirect Evolutionary Analysis. The Economic Journal 116, 1037-1056.

[7] Harsanyi J. C. and R. Selten, 1988. A General Theory of Equilibrium Selection. Princeton, The MIT Press.

[8] Kliemt, H., 1986. The veil of insignificance. European Journal of Political Economy 2 (3), 333-344.

[9] Kirchgässner,G. and W. W. Pommerehne, 1993. Low-cost decisions as a challenge to public choice. Public Choice 77, 107-115.

[10] Lewis, D. K., 1969. Convention. Cambridge, Harvard University Press.

[11] Miller, L. M., 2007. Coordinación y convención: Un análisis teórico y experimental de la noción de convención social, Doctoral Thesis, Universidad Complutense de Madrid.

[12] Nozick, R., 1993. The Nature of Rationality. Princeton, Princeton University Press.

[13] Rydval, O. and A. Ortmann, 2005. Loss avoidance as selection principle: Evidence from simple stag-hunt games. Economic Letters 88, 101-107.

[14] Sacconi, L. and S. Moretti, 2002. Fuzzy Norms, Default Reasoning and Equilibrium Selection in Games under Unforeseen Contingencies and Incomplete Knowledge. Liuc Papers in Ethics, Law and Economics 104, Catteneo University.

[15] Schotter, A., 1981. The Economic Theory of Social Institutions. Cambridge, Cambridge University Press.

[16] Straub, P. G., 1995. Risk Dominance and Coordination Failures in Static Games, The Quarterly Review of Economics and Finance 35 (4), 339-363.

[17] Sugden, R., 1986. The Economics of Rights, Cooperation and Welfare. Oxford, Basil Blackwell.

[18] Ullmann-Margalit, E., 1977. The Emergence of Norms. Oxford, Clarendon Press.

[19] Van Damme, E., 1987. Stability and Perfection of Nash Equilibria. London, Springer.

[20] Van Huyck, J. B. , R. C. Battalio, and R. O. Beil, 1990. Tacit coordination games, strategic uncertainty, and coordination failure. American Economics Review 80 (1), 234-248.

[21] Van Huyck, J. B. , R. C. Battalio, and R. O. Beil, 1991. Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. Quarterly Journal of Economics 106 (3), 885-911.

[22] Van Huyck, J. B., J. P. Cook, and R. C. Battalio, 1997. Adaptive behavior and coordination failure. Journal of Economic Behavior and Organization 32 (4), 483-503.

[23] Young, H. P., 1993. The Evolution of Conventions, Econometrica 61 (1), 57-84.

[24] Young, H. P., 1996. The Economics of Convention. Journal of Economic Perspectives 10 (2), 105-122.

[25] Young, H. P., 1998. Individual Strategy and the Social Structure. An Evolutionary Theory of Institutions. Princeton, Princeton University Press.

Appendix A: Evolutionary analysis

Success table (part 1)

$R(\widetilde{T},T)$		U	MG	T S	T R	T L	I	A
$\widetilde{T}$	U	1+a for 1+d>c & a+b>1	b+1 for 1+d>c & a+b>1	a+c for 1+d>c & a+b>1 & a>c; b+d for 1+d>c & a+b>1 & a<c	a+1 for 1+d>c & a+b>1 & a>(c-d)(1-b); d for 1+d>c & a+b>1 & a<(c-d)(1-b)	d+b for 1+d>c & a+b>1		(1+a+b+d)/2 for 1+d>c & a+b>1
		d for 1+d>c & a+b<1	2 for 1+d>c & a+b<1	1 for 1+d>c & a+b<1 & a>c; 1+d for 1+d>c & a+b<1 & a<c	a+d for 1+d>c & a+b<1 & a>(c-d)(1-b); c for 1+d>c & a+b<1 & a<(c-d)(1-b)	1+d for 1+d>c & a+b<1		(2+d)/2 for 1+d>c & a+b<1
		b for 1+d<c & a+b>1	b for 1+d<c & a+b>1	a for 1+d<c & a+b>1 & a>c; c+b for 1+d<c & a+b>1 & a<c	1+b for 1+d<c & a+b>1 & a>(c-d)(1-b); 1 for 1+d<c & a+b>1 & a<(c-d)(1-b)	b+c for 1+d<c & a+b>1		(a+b+c)/2 for 1+d<c & a+b>1
		c+1 for 1+d<c & a+b<1	1 for 1+d<c & a+b<1	0 for 1+d<c & a+b<1 & a>c; c+1 for 1+d<c & a+b<1 & a<c	b+d for 1+d<c & a+b<1 & a>(c-d)(1-b); c+1 for 1+d<c & a+b<1 & a<(c-d)(1-b)	b+c for 1+d<c & a+b<1		(1+c)/2 for 1+d<c & a+b<1
	MG	a for 1+d>c & a+b>1; c+a for 1+d>c & a+b<1; b for 1+d<c & a+b>1; c+b for 1+d<c & a+b<1	b	a for a>c; b+c for a<c	a for a>(c-d)(1-b); b+c for a<(c-d)(1-b)	c+b	(c-1)+(c-d)+(a,1-b+a)	(a+b+c)/2
	S	1+a for 1+d>c & a>c; 0 for 1+d>c & a<c; a+d for 1+d>c & a+b<1 & a>c; c for 1+d<c & a+b<1 & a<c; b+1 for 1+d<c & a+b>1 & a>c; 1 for 1+d<c & a+b>1 & a<c; b+d for 1+d<c & a+b<1 & a>c; c+1 for 1+d<c & a+b<1 & a<c	b+1 for a>c; 1 for a<c	a+1 for a>c; c+1 for a<c	a+1 for a> (c-d)(1-b) & a>c; 0 for a> (c-d)(1-b) & a<c; b+d for a< (c-d)(1-b) & a>c; c+1 for a< (c-d)(1-b) & a<c	d+b for a>c; 0 for a<c		(1+a+b+d)/2 for a>c; (1+c)/2 for a<c

18

Success table (part 2)

$\hat{\pi}$						
R	1+a for 1+d>c & a+b>1 & a>(c-d)(1-b) d for 1+d>c & a+b>1 & a<(c-d)(1-b) 1 for 1+d>c & a+b<1 & a>(c-d)(1-b) d+1 for 1+d>c & a+b<1 & a<(c-d)(1-b) a for 1+d<c & a+b>1 & a>(c-d)(1-b) c+b for 1+d<c & a+b>1 & a>(c-d)(1-b) 0 for 1+d<c & a+b<1 & a>(c-d)(1-b) c+1 for 1+d<c & a+b<1 & a<(c-d)(1-b)	1+b for a>(c-d)(1-b)	a+1 for a>(c-d)(1-b) & a>c d+b for a>(c-d)(1-b) & a<c	1+a for a>(c-d)(1-b)	d+b for a>(c-d)(1-b)	(1+a+b+d)/2 for a>(c-d)(1-b)
L	0 for 1+d>c & a+b>1 c for 1+d>c & a+b<1 1 for 1+d<c & a+b>1 1 for 1+d<c & a+b<1	1	0 for a>c 1+c for a<c	0 for a>(c-d)(1-b) 1+c for a<(c-d)(1-b)	1+c	(1+c)/2
I	(b)(c-d)+1 1+c-d)+ (1-b)(1-b+a) for 1+d>c & a+b>1 (a)(c-d)/1+c-d)+ (d)(1-b)+(a)(1-b+a) for 1+d>c & a+b<1 (b)(c-d)+1 1+c-d)+ (1-b)(1-b+a) for 1+d<c & a+b>1 (b)(c-d)+1 1+c-d)+ (d)(1-b)+(a)(1-b+a) for 1+d<c & a+b<1	(b)(c-d)+1 1+c-d)+ (1-b)(1-b+a)	(a)(c-d)/1+c-d) + (1-b)(1-b+a) for a>c (b)(c-d)+1 1+c-d)+ (d)(1-b)+(a)(1-b+a) for a<c	(a)(c-d)+1 1+c-d) + (1-b)(1-b+a) for a>(c-d)(1-b) (b)(c-d)+1 1+c-d)+ (d)(1-b)+(a)(1-b+a) for a<(c-d)(1-b)	(b)(c-d)+1 1+c-d)+ (d)(1-b)+(a)(1-b+a)	(c-1)+(c-d) + (a)(1-b+a)
A	(1+a)/2 for 1+d>c & a+b>1 (a+c+d)/2 for 1+d>c & a+b<1 (2+b)/2 for 1+d<c & a+b>1 (1+b+c+d)/2 for 1+d<c & a+b<1	(2+b)/2	(1+a)/2 for a>c (1+b+c+d)/2 for a<c	(1+a)/2 for a>(c-d)(1-b) (1+b+c+d)/2 for a<(c-d)(1-b)	(1+b+c+d)/2	(2+a+c)/2

Appendix B: Experimental instructions

Welcome and thank you for participating in this experiment. Please read the instructions carefully. From now on, any communication with the other participants is forbidden. Otherwise, we would have to exclude you from the experiment and from all payments. If you have any questions, please raise your hand. We will come to you and answer your questions individually.

During the experiment, we use ECU (experimental currency unit) instead of euro. Your experimental income will be calculated in ECU. At the end of the experiment, the ECU you have earned, will be converted to euro (1 ECU = 0.5 euro) and the obtained amount will be paid to you in cash.

In this experiment, you will interact with one other participant, and each of you will play a different role which, to simplify matters, is either A or B. Both participants can choose between two strategies X_A or Y_A (for participant A) and X_B or Y_B (for participant B).

These strategies lead to the following payments:

Payment for participant A	Payment for participant B	If participant A selects	If participant B selects
1	a	X_A	X_B
d	0	X_A	Y_B
0	b	Y_A	X_B
c	1	Y_A	Y_B

Values a, b, c and d will fulfill the following conditions:

$$1 > a > b > 0 \text{ and } 1 > c > 0 > d > -1$$

Whether X_i or Y_i is better for you depends on the decision of your partner in this round, e.g., on which alternative he or she selects. You should therefore be interested in collectively arranging a strategy constellation $(X_A, X_B), (X_A, Y_B), (Y_A, X_B)$, or (Y_A, Y_B).

To reach such an agreement, we propose principles by which you can select one of the four strategy constellations. (These principles will be explained to you on the next page.)

First, you have to rank principles. You must not (but you may) rank all principles. It suffices to indicate your priority. Then, knowing both ranking orders (yours and that of the participant with whom you interact), you can choose either strategy X or strategy Y. In each round, you will learn which role you play (A or B) as well as the values of the parameters a, b, c, and d. Then you play another nine rounds of the game (i.e., in total you play ten rounds). At the begining of each round, the values of the parameters a, b, c, and d and your role (A or B) will change. Your selected strategy rank order remains the same.

After these ten rounds, you will be matched randomly with a new partner with whom you interact for another ten rounds. You will keep your strategy

rank order and will also learn the strategy rank order of your partner. In each round, you will learn again which role you play (A or B) as well as the values of the parameters a, b, c, and d. At the end of these 20 rounds, we will inform you about the values of the parameters a, b, c, and d in the individual rounds, whether you have been participant A or B in the individual rounds, and about your payoff in each round.

To compensate possible losses, each participant receives a basic amount of 4 ECU (2 ruros). Should you finish the experiment with a negative balance, we will ask you to compensate this by doing some odd jobs (for 32 ECU per hour) at our institute. If you don't agree, please raise your hand. You may then leave the experiment.

We will now present to you the principles for selecting one of the four strategy constellations:

A

In even rounds, we play X and in uneven rounds Y so that both participants rely either on their X (resp. Y) choice.

G

We use the strategy which is better for the other participants if he or she plays both strategies with the same probability, i.e., participant A should play Y_A and participant B should play X_B.

I

We use both strategies with such positive probabilities that it will not matter for the other participants what he or she selects.

L

To avoid losses, we play Y, i.e., Y_A, Y_B.

M

We use the principle which has the better of the worst possible results, i.e., participant A should always play Y_A and participant B always X_B.

R

We use the strategy constellation where each participant reacts optimally to the choice of the other, and the product of the corresponding payoff increases resulting from a strategy change is maximal.

For participant A the payoff increases by $1 - 0 = 1$ for X_B and $c - d$ for Y_B. For participant B the increase is $a - 0 = a$ for X_A and $1 - b$ for Y_A. As both participants react optimally to the choice of the respective other only for (X_A, X_B) and (Y_A, Y_B), they should use (X_A, X_B) if $a > (c - d)(1 - b)$ and (Y_A, Y_B) if $a < (c - d)(1 - b)$.

S

We use the strategy constellation for which the payoff sum is maximal, i.e. participants play (X_A, X_B) if $a > c$ and (Y_A, Y_B) if $a < c$.

U

We use the strategy which is better if the other uses both strategies with the same probability, i.e. participant A should play X_A if $1 + d > c$ and Y_A if $1 + d < c$. Participant B should play X_B if $a + b > 1$ and Y_B if $a + b < 1$.

Round	Gains		
	Strategy	Gains A	Gains B
1	X_A, X_B	1	0,5
	X_A, Y_B	-0,5	0
	Y_A, X_B	0	0,2
	Y_A, Y_B	0,4	1

Round	Gains		
	Strategy	Gains A	Gains B
2	X_A, X_B	1	0,8
	X_A, Y_B	-0,1	0
	Y_A, X_B	0	0,1
	Y_A, Y_B	0,7	1

Round	Gains		
	Strategy	Gains A	Gains B
3	X_A, X_B	1	0,9
	X_A, Y_B	-0,3	0
	Y_A, X_B	0	0,8
	Y_A, Y_B	0,6	1

Round	Gains		
	Strategy	Gains A	Gains B
4	X_A, X_B	1	0,7
	X_A, Y_B	-0,2	0
	Y_A, X_B	0	0,6
	Y_A, Y_B	0,5	1

Round	Gains		
	Strategy	Gains A	Gains B
5	X_A, X_B	1	0,6
	X_A, Y_B	-0,4	0
	Y_A, X_B	0	0,5
	Y_A, Y_B	0,7	1

Round	Gains		
	Strategy	Gains A	Gains B
6	X_A, X_B	1	0,3
	X_A, Y_B	-0,8	0
	Y_A, X_B	0	0,2
	Y_A, Y_B	0,6	1

Round	Gains		
	Strategy	Gains A	Gains B
7	X_A, X_B	1	0,7
	X_A, Y_B	-0,6	0
	Y_A, X_B	0	0,2
	Y_A, Y_B	0,9	1

Round	Gains		
	Strategy	Gains A	Gains B
8	X_A, X_B	1	0,5
	X_A, Y_B	-0,1	0
	Y_A, X_B	0	0,2
	Y_A, Y_B	0,5	1

Round	Gains		
	Strategy	Gains A	Gains B
9	X_A, X_B	1	0,2
	X_A, Y_B	-0,2	0
	Y_A, X_B	0	0,1
	Y_A, Y_B	0,5	1

Round	Gains		
	Strategy	Gains A	Gains B
10	X_A, X_B	1	0,3
	X_A, Y_B	-0,2	0
	Y_A, X_B	0	0,2
	Y_A, Y_B	0,6	1

Appendix C: Computer screens

Ranking screen

Decision-making screen