



**Perceiving strategic
environments:
An experimental study of
learning under minimal
information**

Andreas Nicklisch





Perceiving strategic environments: An experimental study of learning under minimal information

Andreas Nicklisch

June 2006

Perceiving strategic environments: An experimental study of learning under minimal information*

Andreas Nicklisch[†]

Abstract

We present the results of an experiment on learning with minimal information. Particularly, subjects are only provided with feedback about their own payoff from the last period of the game being played, but not with information about the structure of the game. We compare the empirical structure of the decision algorithm for this setting with the empirical structure of algorithms for subjects who receive sufficient information to learn the game. The laboratory data show that, depending on the information setting, players adjust their strategy choice differently. The structure of the decision algorithm for subjects operating with minimal information indicates myopic responses to success, while the structure for sufficiently informed players is more complex. As a consequence, sufficiently informed players outperform players who have minimal information in a simple coordination game. Yet, if the structure of the game changes, readjustment is more successful for the players operating with minimal information.

Keywords: Experimental economics, learning, minimal social situation, myopia

[*APA*] 2340, 2343

[*JEL*] D83, D84

*This article is based on an earlier working paper, "Perceiving strategic environments: An experimental study of strategy formation and transfer". The author would like to thank Yoella Bereby-Meyer, Arndt Bröder, Christof Engel, Axel Ockenfels, and Isabelle Schnabel for their support.

[†]Max Planck Institute for Research on Collective Goods, Bonn, Germany; email nicklisch@coll.mpg.de

1 Introduction

During the last decade, there has been an increasing economic literature on individual learning theories (e.g., Crawford, 1995, Mookherjee & Sopher, 1997, Fudenberg & Levine, 1998, Huck et al., 1999). Those theories apply algorithms to predict individual adaptation processes. Experimental data show that algorithms that base decisions on a rich set of information, for instance fictitious play, poorly describe human behavior.¹ Rather, algorithms of limited cognitive sophistication are more useful descriptions of real decision making (e.g., Roth & Erev, 1995, Erev & Roth, 1998, Camerer & Ho, 1999). Among all pieces of information, one's own payoffs appear to be the most relevant.² Even for fairly complex environments, Nagel & Vriend (1999) show that subjects' behavior can best be described by the simple rule of directional learning relating individual actions and individual payoffs. The experimental evidence suggests that subjects use the provided information quite selectively; unnecessary information is neglected.

We want to pursue the issue in the opposite direction. By and large, the literature on individual learning is concerned with which information is important for the decision algorithm. The objective of our paper is to investigate how the structure of a particular decision algorithm responds to the information received about the environment. Especially, we are interested in exploring how subjects behave under conditions of minimal information, that is, if they do not know the payoff matrix of the game, and if their own payoffs are the only feedback players receive, and how the structure changes if information is enriched. For this purpose, the structure of the decision algorithm both for players with minimal information and for players with additional information is estimated and compared. There are only a few studies dealing with learning under conditions of minimal information (e.g., Arickx & van Avermaet, 1981, Mitropoulos, 2001). The general results suggest that the restricted feedback leads to more inertia in players' actions than they exhibit in settings with a richer feedback (Mookherjee & Sopher, 1994, van Huyck et al., 2005). In oligopoly experiments, inertia of less informed players leads to less competitive prices/quantities (Huck et al., 1999); later studies, however, cannot confirm this result for actions in continuous time (Friedman et al., 2004).

For the estimation, we use the experience-weighted attraction learning model (hereafter EWA) by Camerer & Ho (1999). The EWA model

¹Fudenberg & Levine (1998) extensively discuss fictitious play.

²Strategy choice obeys the "law of effect" (e.g., Herrnstein, 1970) as strategies which have been successful are played more frequently than those which have been less successful.

provides us with a framework assessing the explanatory power of simple decision algorithms, like reinforcement learning, and directional learning. Prior to playing, subjects are not informed about the structure of the game. However, some players receive sufficient information in order to learn the structure of the game, while other do not. We focus in our experiment on two simple 2x2 games, the *mutual fate game* (hereafter MFC) and the *fate control/behavior control game* (hereafter FCBC). We will introduce the two games in Section 2. We chose these two games since successful coordination in the games requires that the decision algorithm has a specific structure. In Section 3, we discuss the EWA model, along with the structure it needs for the MFC and the FCBC. Section 4 reports the experimental data. Section 5 discusses the results of our experiment.

2 The games and the experimental design

In the experiment, all subjects are informed that they will be playing a simple 2x2 game, that is, two players, two alternative actions, that involve only two possible payoff levels, either 0 or 1. Consequently, in principle, players face 2^8 possible games.³ Eliminating all the trivial games⁴ and all games with a weakly dominated action, four different types of games remain: the matching pennies games (where one player wins one point when he chooses the same action as the opponent, while the other wins when he chooses the opposite action of the other), a simple coordination game (where both players win a point when they both choose the same action), MFC, and FCBC (Mitropoulos, 2001). Psychologists are more familiar with MFC since it characterizes what psychologists define as a minimal social situation (Sidowski et al., 1956, Kelley et al., 1962). Here, one of at least two persons affects the payoff of at least one other player by choosing an action – possibly even without knowing of this effect. Crucially for MFC, the individual payoff of one player depends exclusively on the actions by the other player, but is not influenced by the action that one chooses. The resulting bimatrix of the game is shown in Table 1.

Of course, any mixed strategy vector is a weakly dominated Nash equilibrium. One can think of the following illustrative example. Consider two train compartments, denoted as left and right. Both compart-

³This setting is different from that in experimental detecting games, which test the ability of subjects to investigate the type of counterpart they have, while knowing what type they themselves are (e.g., Cox, Shachat & Walker, 2001 and Oechssler & Schipper, 2003). We do not have any prior information about any type of player.

⁴Games where both players receive the payoff / do not receive any payoff, irrespective of the action chosen.

	player 2	a	b
player 1			
a		1	1
	a	1	0
b		0	0
	b	1	0

Table 1: *The payoff matrix for the mutual fate control game*

ments are equipped with a two-scaled air-conditioning system. However, the system control in the right compartment controls the system in the left compartment, and vice versa. Let us assume that travelers are not allowed to leave their compartments. The interesting question now refers to how travelers in the left compartment signal the travelers in the right compartment their desire for low/high temperatures by use of the system control in their compartment. Simultaneously, travelers in the right compartment try to signal their needs using their system control.

The second game we use for our laboratory experiment differs only slightly from the MFC. The fate control/behavior control game was first introduced by Thibaut & Kelley (1959) and Rabinowitz et al. (1966). Whereas in the MFC, any mixed strategy vector is a weak Nash equilibrium for both players, in the FCBC, for one player there is a best response to the action of the other player. The relationship among the players can be characterized as follows: player 1's action determines the payoff of player 2, regardless the action chosen by player 2. Therefore, any mixed strategy vector is a best response by player 2. However, player 1 has a distinct best answer to the action chosen by his/her counterpart, that is, player 1's payoff depends on player 1's *and* player 2's action. The resulting bimatrix for the payoffs of the FCBC is shown in Table 2.

For p , indicating the probability of player 1 choosing a , and q , the probability of player 2 choosing a , any q for $0 \leq q \leq 1$ belongs to an equilibrium vector (p, q) . However, for player 1 it is possible to specify three optimal responses; thus, we have the three types of equilibrium vectors (p, q) in the FCBC: $(\frac{1}{2}, \frac{1}{2})$, $(0, q')$, and $(1, q'')$ for all q' and q'' , respectively, such that $0 \leq q' < \frac{1}{2} < q'' \leq 1$.

The experiment consists of 4 phases. Pairs of players remain in the same group throughout the entire experiment. Participants first face two phases of 30 periods each of the MFC. Within each phase, the structure of the payoff matrix does not change. Between phases, we permute the payoff matrix of the MFC and inform players that the structure of the game has changed. After two phases of the MFC, we permu-

	player 2	a	b
player 1			
a		1	1
b		0	0
		0	1

Table 2: *The payoff matrix for the fate control/behavior control game*

tate the structure of the payoff matrix such that subjects face a FCBC. Again, the structural change is announced, and subjects play 30 periods. Thereafter, we permute the payoff structure such that players face another FCBC. After announcing the structural change, subjects play for another 30 periods.

The complexity differs substantially if the games are played under the minimal information setting or with sufficient information, allowing players to understand the structure of the game. In particular, we will observe decisions of players with the following degrees of information about the games:

- (*The minimal information setting*) Both players are informed that they will be playing a symmetric two-person game⁵ with two choices and a payoff of either 0 or 1. As feedback, they receive only their own payoff. We refer to these subjects as the low informed players (hereafter LI players).
- (*The rich information setting*) In addition to the information available in the LI condition, both players observe the choice of their counterparts in the same period. We refer to these subjects as the rich informed players (hereafter RI players).

Note that LI players cannot understand the structure of the game. They may observe that there is little relation between their own action and their own payoffs, but they cannot learn that the action of the counterpart exclusively determines their own payoffs. In contrast, RI players can learn the structure of the games by disentangling the payoffs for all four cells of the games. In addition to these two treatment conditions, one may argue that additional incentives facilitate the understanding of the game's structure. Subjects may search more systematically if finan-

⁵Of course, we informed players that symmetry is not retained in phases two, three and four.

cial incentives vary across the experiment. Therefore, we introduce the following treatment condition:

- (*The rich information plus setting*) Both players receive the same information as RI players. However, their payoff is either 0 or 1 in the first phase of the experiment, 0 or 2 in the second phase, 0 or 3 in the third phase, and 0 or 4 in the fourth phase. We refer to these subjects as the rich informed “plus” players (hereafter RI+ players).

Finally, we want to disentangle the behavior of players who received minimal information and the influence due to having previously played the MFC. For this purpose, we consider minimally informed players who only participated in the FCBC.

- (*The control setting*) Both players receive the same information as LI players, but play only two phases of the FCBC. We refer to these subjects as the control players (hereafter CO players).

In the experiment, both subjects in each group are either LI players, RI players, RI+ players, or CO players. At the end of each phase, participants are asked to describe the payoff mechanism in an open question. Furthermore, in order to estimate subjects’ short-term memory capacity after the experiment, subjects are asked to participate in a Wechsler digit span test (Wechsler, 1945). For one second we present random numbers with increasing digits, which have to be repeated by subjects after the number disappeared. We use the number of the last digits correctly reported as an approximation of the short-term memory capacity of the subjects. We introduce both the open question and the memory test to get more insights about the relation between memory capacity, information levels, and the correct understanding of the games’ structure. Yet, we have to mention that no monetary incentives are provided for either task.

3 Learning and the minimal information setting

It seems that the MFC and the FCBC are rather simple. However, if played under the minimal information setting, the situation changes dramatically. We argue that the MFC and the FCBC are complex games, since, referring to measurements of the cognitive complexity of games, the fact that the players’ own payoffs depend exclusively on the action the counterpart’s choice is puzzling. Schelling (1960) provides the basic distinction between pure-motive and mixed-motive games. Pure-motive games are easy to grasp since they are characterized by a clear-cut correlation between the two players’ preferences for their actions; yet, the

non-correlated structure of preferences of mixed-motive games is inherently difficult to understand. More formalized, Devetag and Warlien (2005) develop a classification of relational complexity of two-person normal form games that is based on the players' preferences for the order of payoff cells. They show, in an experiment where subjects first have to develop a representation of games of different complexity, that the easier subjects can derive the preferable order of cells by their counterpart from their own order of cells, the better they understand the structure of the game. The better subjects understand the structure, the better they perform while playing the game. In accordance with earlier studies (Nyarko & Schotter, 2000), Devetag and Warlien find that subjects who misrepresent the games behave in a way that is consistent with these representations.

Certainly, the MFC and the FCBC are neither pure-motive games nor games of low relational complexity. The important difficulty arises from the fact that, for both players in the MFC and for one player in the FCBC, any (mixed) action choice is an element of a Nash equilibrium. Previous theoretical studies by Jordan (1991, 1995) demonstrated that, given learning processes, in finite-player, finite-action games played under minimal information setting, behavior converges to Nash equilibria. The interesting – but more demanding – question is whether subjects can coordinate, after repeating the game, on the particular equilibrium that yields the Pareto efficient outcome, that is, action “*a*” for the bimatrices in Figure 1 and Figure 2. Hereafter, we denote this coordination as Pareto efficient coordination, PEC. In previous experiments for the MFC (Mitropoulos, 2001), 54% of subjects choose the PEC after 10 periods of play. However, the frequency remains far below the rate that one expects for a convergence towards the PEC. After 100 periods, Mitropoulos finds only 66% of subjects choosing the PEC. This result confirms earlier experiments reported by psychologists (e.g., Rabinowitz et al., 1966, Arickx & van Avermaet, 1981). However, in those experiments, subjects typically did not know whether they were playing a two-person game, and incentives were non-monetary.

Apparently, the difficulty of successfully coordinating varies in accordance with the degree of information about the game. Moreover, it seems plausible to assume that the degree of information crucially changes the way players make their decisions. Hence, we should see differences for the estimated decision algorithms depending on the degree of information. For all estimations, we use a simplified version of the EWA learning model proposed by Camerer and Ho (1999), since this model is a very general representation of several learning rules. The EWA model refers to the attractiveness of actions which are updated based on pay-

off experience. Unlike the standard reinforcement learning (e.g., Roth & Erev, 1995), EWA also considers the hypothetical reinforcement of actions that were not chosen (according to the payoff they would have yielded). In particular, we define, for player i , the attraction of a certain action a_i^j to be chosen in period $t + 1$, $A_i^j(t + 1)$ as

$$A_i^j(t + 1) = \rho_1 A_i^j(t) + \rho_2 \pi_i(a_i^j, a_{-i}(t)) + \rho_3 I(a_i^j) \pi_i(a_i^j, a_{-i}(t)), \quad (1)$$

where $\pi_i(a_i^j, a_{-i}(t))$ denotes player i 's payoff, yielded by i 's choice of action a_i^j , given the choice $a_{-i}(t)$ by all players other than i in period t . Furthermore $I(a_i^j)$ equals 1 if player i actually chose a_i^j in period t ; it is otherwise 0.⁶ Thus, the coefficient ρ_3 reflects the difference in the influence between a payoff yielded and a payoff that would have been yielded. On the other hand, the coefficient ρ_1 indicates the autoregressive influence of earlier experience on the current attractiveness of a certain action, that is, the non-myopical influence for the decisions.

Depending on the parametrization of $A_i^j(t + 1)$, the EWA allows us to characterize different learning rules. A very common type of simple learning is reinforcement learning. For reinforcement learning, only actually yielded payoffs influence attractiveness, so that we expect $\rho_1 \neq 0$, $\rho_3 > 0$, but $\rho_2 = 0$. However, Mitropoulos (2001) shows for the MFC that reinforcement learning fails to provide PEC, and it describes the behavior of LI players rather poorly. Therefore, we need a different approach to describe the behavior of players who reach PEC. This approach is rather simple and follows the basic structure of learning direction theory (Selten & Stöcker, 1986). The theory posits that players change their choices in accordance with a qualitative picture of the world. Whenever they get a negative feedback (no payoff per period), they change the direction of their action. Specifically for the MFC, Mitropoulos (2001) offers the win-stay/lose-change adaptation scheme (hereafter ws/lc), which leads, for the MFC, to coordination on the PEC.⁷ Note that the ws/lc decision rule leads to the restriction of $\rho_2 > 0$, while $\rho_1 = \rho_3 = 0$. Thus, players learn to act myopically. Figure 1 shows the evolution of action if both players follow ws/lc. Let us assume that players find themselves

⁶The original EWA model provides a richer set of parameters, a depreciation rate σ_1 that measures the fractional impact of previous experience, and a discount factor σ_2 that depreciates previous attraction. The factors depend on the amount of previous experience in order to capture the “power-law of practice”, i.e., the influence of new experience decreases with experience. We will not focus on these parameters, as they are not essential for our further argumentation. Nonetheless, they can be reproduced from our set of parameters by $\sigma_1(t) = \frac{1-\rho_2-\rho_3}{(\rho_2+\rho_3)t}$ and $\sigma_2(t) = \frac{\rho_1}{(\rho_2+\rho_3)t}$, with t denoting the amount of previous experience.

⁷This rule shows some similarities to the class of search-and-select adaptation rules proposed by Conlisk (1993).

period	t	$t + 1$	$t + 2$	$t + 3$	
player 1	$b(1)$	$b(0)$	$a(1)$	$a(1)$...
player 2	$a(0)$	$b(0)$	$a(1)$	$a(1)$...

Figure 1: *The ws/lc adaptation for the MFC (payoffs in parenthesis)*

in the constellation of period t . Then, ws/lc leads to PEC within the next two periods. If, however, players find themselves in the constellation of period $t + 1$, PEC is reached in the next period. Finally, if both players follow ws/lc, their actions do not change thereafter, that is, in period $t + 2$ and later periods. It is important to stress that all players, LI, RI, RI+, and CO players, receive such a degree of information that they can develop a reinforcement learning algorithm. However, ws/lc requires an understanding of the basic structure of the MFC; otherwise the prediction of $\rho_2 > 0$ seems counterintuitive. Thus, we test the data of RI and RI+ players for the ws/lc. For reasons of completeness, we also retain ρ_2 in the estimation for LI players, but do not assume $\rho_2 > 0$ here. Previous experimental evidence by Mitropoulos (2001) suggests that subjects acquire a representation of the game that is much more complex. Particularly, he reports frequent individual patterns such as an alternation in their choices of actions or in the sequences of certain changes in actions. Empirical behavior must therefore be modelled such that $\rho_2 > 0$, $\rho_3 > 0$, and, especially, $\rho_1 \neq 0$, for example, that alternating changes of actions would result in $\rho_1 < 0$.

More abstractly, ws/lc learning refers to the most simple algorithm with the least number of significant coefficients of the EWA model. One can argue that the rich information setting – and even more, the rich information plus setting – facilitates sufficient information about MFC, and, therefore, promotes decisions to follow the simple ws/lc rule. Consequently, one expects that players in the minimal information setting will be free to interpret the MFC in various ways since they cannot understand the game. Thus, we can assume that the estimated structure of LI players’ decision algorithms will be more complex than those of RI and RI+ players, since LI players have no chance to detect the structure of the game.

On the other hand, one can claim that more information facilitates decision algorithms with more complex structures. This argument rests on the idea that the less players know about the game, the less can they (mis)interpret its structure. Thus, we have to test whether the structure of LI players’ decision algorithm is less complex than those of RI and RI+ players.

period	t	$t + 1$	$t + 2$	$t + 3$	
player 1	$b(0)$	$a(0)$	$b(1)$	$b(0)$...
player 2	$a(0)$	$b(1)$	$b(0)$	$a(0)$...

Figure 2: *The ws/lc cycle moves in the FCBC (payoffs in parenthesis)*

A major difficulty arises for subjects who apply the successfully learned ws/lc strategy for the PEC in the FCBC. Unless they start with PEC by pure chance, ws/lc creates cycle moves (see Rabinowitz et al., 1966). Figure 2 schematically shows the cycle moves for the FCBC according to ws/lc. Suppose players find themselves in the constellation of period t . Then ws/lc leads to a change in action for both players so that they end up in the constellation of period $t + 1$. Departing from period $t + 1$, player 1 changes his action, while player 2 does not; this leads to the constellation of period $t + 2$. Therefore, player 1 retains in the same course of action, whereas player 2 changes his course of action in $t + 3$. As a result, the evolution of action ends again in a constellation equivalent to t , and the PEC is not reached on this cycle.

Thus, experienced players, that is, subjects who learned to adapt their actions according to ws/lc, have to unlearn myopia if they are to be successful in the subsequent FCBC game. Note that no unlearning is required if players, under the minimal information setting, acquired a more complex, suboptimal decision algorithm.

4 Experimental results

4.1 The mutual fate control game

Experiments took place at the computer laboratory of the Max Planck Institute for Economics in Jena, Germany, and at the computer laboratory of the University of Bonn, Germany. In total, 160 mostly undergraduate students participated in the computerized experiment, using zTree software (Fischbacher, 1999). Subjects were recruited using ORSEE software (Greiner, 2004). Thus, we observe, in sum, 20 pairs for each condition. Subjects needed approximately 60 minutes for the entire experiment (35 minutes in the control condition).⁸ During the experiment, participants received payoffs in experimental currency units, which were converted afterwards to euros. Average earnings were quite surprising. Including a 2.50 euro show-up fee, on average, LI players earned 8.30

⁸Due to timing problems of participants, we had to dismiss 29 subjects across all treatment conditions before they performed in the digit span test. Therefore, we have no approximation of their short-term memory capacity.

euros (standard deviation 1.26), RI players only slightly more, at 8.70 euros (sd 1.33), while RI+ players were paid 16.30 euros (sd 1.99), and CO players 5.50 euros (sd 2.13).

It is not surprising that the comparison of the first phase frequencies of PEC for LI players and RI players suggests that the richer information level facilitates coordination. Throughout the first 10 periods of the first phase, 55% of the RI players chose the PEC. For the last 10 periods, this number increased to 67%. On the other hand, LI players started with 49.5% for the first ten periods and ended with 60.8%.⁹ The latter results correspond roughly to findings that have been previously reported (e.g., Mitropoulos, 2001). Figure 3(*i*) shows the development for 5 period averages of the PEC across the first phase. However, the result changes dramatically if we consider the coordination frequencies for RI+ players. These players chose the PEC in 46% of the cases across the first ten periods and in 56.8% across the last ten periods. Note that there is a significant difference between the averages of RI and RI+ players for the first ten periods and the last ten periods, while the other frequencies do not differ significantly.¹⁰ It seems that the additional incentive for RI+ players, that is, the prospect of earning significantly more in later phases of the experiment, hampers coordination in the first phase. We will discuss this point later in more depth.

Quite surprisingly, in the second phase of the MFC, the LI players outperform RI players in terms of the frequency of PEC. Figure 3(*ii*) shows the development for the 5 period averages of the PEC across the second phase. In the last 10 periods, the frequency of LI players significantly exceeds that of RI players (59% for RI and 67.3% for LI players),¹¹ although subjects under both conditions start with an insignificant difference in the frequency of PEC in the first 10 periods (53% for RI and 55.5% for LI players). Moreover, the averages for RI+ players do not progress across the entire second phase. For the first 10 periods, 47% of players chose the PEC. This rate remained at 47% for the last 10 periods. Both rates are significantly lower than the corresponding rates for the RI and the LI players.¹²

Within the last ten periods of the first phase, the frequency of RI pairs who successfully coordinated on the Pareto efficient cell (52%) exceeded the frequency of LI pairs (44.5%) and the frequency of RI+ pairs

⁹Applying two-sided t-tests, all differences differ significantly for $p < 0.05$, except for the comparison of the first ten periods for RI and RI+ players.

¹⁰For the former comparisons, $p < 0.05$ for two-sided t-tests; for the latter comparisons, $p > 0.05$ for two-sided t-tests.

¹¹For the former comparison, $p = 0.49$ for a two-sided t-test; for the latter comparison, $p = 0.016$ for a two-sided t-test.

¹²For all comparisons, $p < 0.05$ for one-sided t-tests.

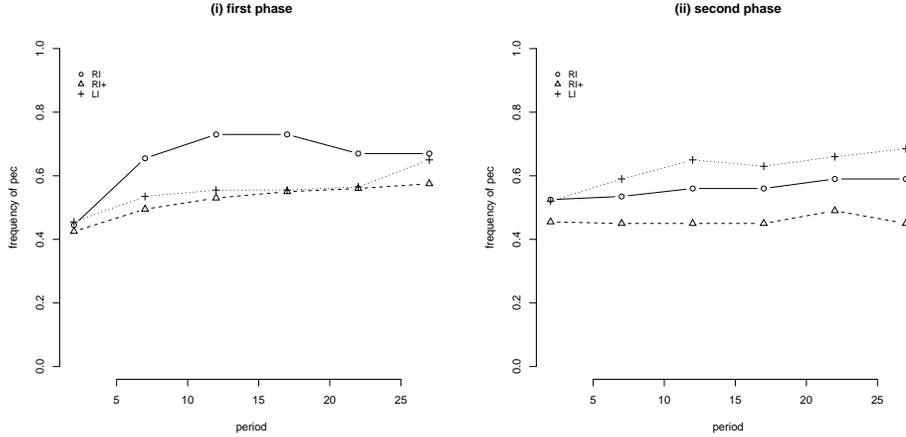


Figure 3: Average frequency of PEC choice in (i) the first phase MFC, (ii) the second phase MFC

(34.5%). In contrast, within the last ten periods of the second phase, 56% of LI pairs coordinated on the Pareto efficient cell, while 47.5% of RI pairs and only 28% of RI+ pairs chose mutual PEC. Moreover, among those LI and RI players who reached PEC in the first phase (second phase), approximately $\frac{7}{10}$ ($\frac{4}{5}$) mutually chose PEC. However, among those RI+ players who reached PEC in the first phase (the second phase), only $\frac{3}{5}$ ($\frac{3}{5}$) mutually chose PEC.

The important question is why RI players perform so poorly in the second phase and why RI+ players perform so poorly throughout both phases of the MFC. Figure 4 offers important insights into both questions. The horizontal axis of Figure 4(i) shows the number of cells of the MFC within which players found themselves during the entire first phase of the experiment, while the vertical axis provides the cumulative frequency of players who experienced, at most, this number of cells. Thus, the figure indicates the proportion of players who experienced all cells of the game. This frequency is less informative for LI players since they cannot disentangle cells and, therefore, cannot understand the structure of the game. Yet, in order to understand the structure of the MFC, RI players and RI+ players have to observe all four cells. For the first phase, one can see that more than 60% of RI players could not understand the structure of the MFC since they did not observe all four cells. On the other hand, the majority of RI+ players observed all four cells of the game. Only 35% of these players were not able to understand the game's structure when they observed three or fewer cells. As shown in Figure 4(ii), the situation becomes even worse in the second phase. More

than 70% of RI players did not observe all four cells. In contrast, the large majority of RI+ players still explored all four cells. It seems that the additional incentive the RI+ players had increasing their prospective earnings led these players to conduct a very systematic search in the first and second phase of the experiment. That the payoff structure of a game crucially influences behavior has been well documented in previous experiments (Merlo & Schotter, 1999).

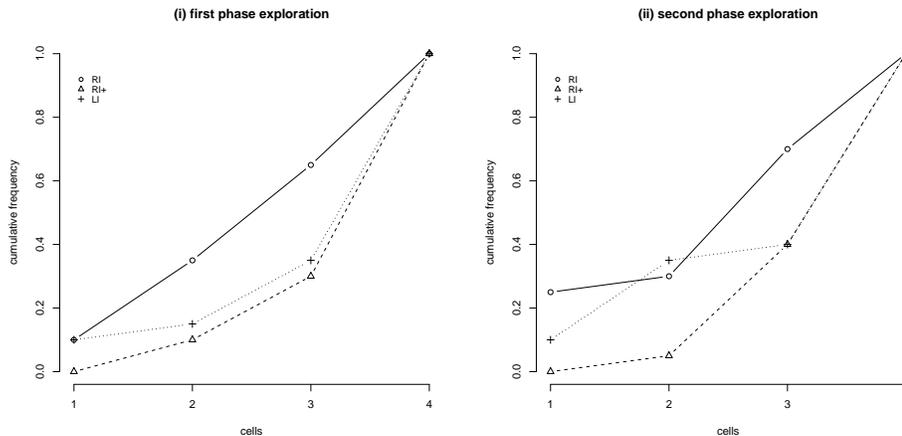


Figure 4: Cells explored in (i) the first phase MFC, (ii) the second phase MFC

Does less information facilitate successful coordination in the MFC? LI players outperform RI players in the second phase of the experiment; RI players outperform RI+ players, that is, players who explored the structure of the game most systematically, in both phases of the MFC. It seems that better informed players fall victim to a curse of knowledge. The analysis of the answers to the open questions for the first and the second phase illustrates, for all treatment conditions, that a majority of players hypothesize a direct relation between their own actions and their own payoffs. This result is not surprising for LI players since they do not receive any information about their counterparts. In particular, 67.5% of the statements after the first phase (62.5% of statements after the second phase) describe a pure coordination game, that is, a relation between individual actions and individual payoffs. Yet, the statements of RI players and those under the RI+ condition demonstrate that their additional information does not prevent them from representing the game “egocentrically”. More specifically, after the first phase, 72% of the statements (after the second phase, 77.5% of the statements) of RI players and 72% of the statements after the first phase (67.5% of

the statements after the second phase) of RI+ players refer to a direct relation between their own actions and their own payoffs. This result follows the robust findings of the psychological literature about the illusion of control; it has been well demonstrated that subjects persistently believe that they can control or at least influence their outcomes even in situations in which this is not true (e.g., Langer, 1975).

Thus, an egocentrically biased representation of the game deceases the performance in the MFC. One could argue that players simply apply some mixed strategies of punishments and rewards.¹³ Yet, applying the Wald-Wolfowitz one-sample run tests, we can analyze whether the number of sequences in which the same action has been played corresponds with the expected number of sequences under a mixture of strategies (for details see Mookherjee & Sopher, 1994). We find that a lot of players show too much inertia for the mixed strategy hypothesis to have explanatory power. For the first phase, we have to reject the mixed strategy hypothesis for 18 LI players, 25 RI players, and 18 RI+ players (out of 40 players within each treatment condition).¹⁴ For the second phase, we reject the mixed strategy hypothesis for 21 LI players, 22 RI players, and 16 RI+ players.¹⁵

Based on this observation, it seems likely that the estimation of decision algorithms indicates that the structure of these algorithms differs quite substantially across treatment conditions. For the approximation of the decision algorithm, we run an individual fixed effect logit regression on the attractiveness of action b ; thus, we estimate the attractiveness in the range $[0, 1]$, with the latent model as defined in equation (1). Choices from phase two are indicated by the dummy variable *two*. Thus, the interaction of coefficients with *two* indicates the learning across phases, for example, $\rho_1 \times two$ tests whether there is a difference in the influence of $A_i^j(t-1)$ between phases one and two. $\rho_2 \times two$ and $\rho_3 \times two$ test the other interaction effects. We set the ex-ante attractiveness of b prior to the first periods to 0.5. As in the original EWA model by Camerer and Ho (1999), we restrict our analysis to a simple autoregressive process with one period lag length ($AR(1)$). Asterisks indicate levels of significance.¹⁶ We report the number of (clustered) observations (nobs); goodness of fit is reported by the log-likelihood (logLik) and the Akaike information criterion (AIC). Table 3 reports the result of our

¹³Of course, this assumes that players know which action punishes or rewards their counterpart.

¹⁴ $p > 0.05$, Wald-Wolfowitz one-sample run tests.

¹⁵ $p > 0.05$, Wald-Wolfowitz one-sample run tests.

¹⁶* significant at a 10% level, ** significant at a 5% level, and *** significant at a 1% level.

	LI	RI	RI+
ρ_1	-0.151	0.243*	-0.009
ρ_2	0.331***	0.542***	0.053
ρ_3	0.468***	0.34**	0.117
$\rho_1 \times two$	0.156	-0.02	-0.056
$\rho_2 \times two$	-0.2*	0.124	0.012
$\rho_3 \times two$	0.156	0.14	0.052
nobs	40	40	40
logLik	-4158	-4427	-4093
AIC	8336	8874	8205

Table 3: *Coefficients of the attractiveness of b in the MFC*

estimations.

The results for the estimations of the coefficients show that the structure of the decision algorithms differ quite substantially across treatment conditions. We failed to characterize the decisions algorithms of RI+ players in the framework of the EWA. Considering the poor performance of RI+ players throughout the two phases of the MFC, it seems likely that they develop very individual strategies and, therefore, fail to coordinate their actions Pareto-efficiently. Unlike RI+ players, we can estimate the decision algorithm for RI players in the framework of the EWA. The significant ρ_1 indicates an autoregressive element in their decision algorithms. Yet, they fail to learn the ws/lc strategy, which predicts $\rho_2 > 0$ and $\rho_3 = 0$. Rather, we find a mixed form of reinforcement learning, for example, $\rho_3 > 0$, and belief learning, for example, $\rho_2 > 0$. It is important to stress that there are not any significant interaction effects for the second phase. Thus RI players rarely change the structure of their decision algorithm from phase one to phase two. The estimated decision algorithm for LI players shows similar patterns for ρ_2 and ρ_3 . However, for LI players, we find some learning across phases, since we have a significant $\rho_2 \times two$. Quite importantly, there is an insignificant coefficient ρ_1 for the first and second phases in which LI players indicate myopic responses. Therefore, we have to stress that the players who have the least information about the game, that is, LI players, develop the least complex decision algorithm and are able to update the algorithm across phases. RI players who develop more complex decision algorithms do not learn across phases. Finally, RI+ players, for example, players whose majority systematically explores all cells of the

game, develop very individual decision algorithms which we could not identify within the EWA.

The short-term memory-capacity approximated by the results of the Wechsler digit span tests exerts very little influence on the performance in the game. There are not any significant correlations between individual Wechsler digit span test result and the achieved points in any of the conditions or phases. If we control for the influence of the result obtained in the Wechsler test by the other player, or the sum of the results obtained by both players, there is not any significant correlation with the points yielded in the games.¹⁷ The result suggests that the ability to precisely recall a larger number of periods does not facilitate PEC.

4.2 The fate control / behavior control game

The analysis of the average frequency of the PEC for the third phase of the experiment indicates that previous experience playing two phases of the MFC under the minimal information setting has a clear influence on the performance in the FCBC. Throughout the first 10 periods of the third phase, the frequency of PEC for LI players is significantly lower than for CO players, whereas the frequencies for RI players and RI+ players differ insignificantly from the frequency for CO players. Specially, 68% of CO players choose the PEC, 61.5% of RI players, 60% of RI+ players, and 52.5% of LI players.¹⁸ For the last 10 periods, the frequency at which CO players choose PEC increases to 77%, while 74.5% of RI players, 66.5% of the RI+ players, but only 54% of the LI players choose the PEC. The frequency of PEC for CO players significantly exceeds the frequencies for LI players and RI+ players,¹⁹ while there is no significant difference between the frequency for RI players and CO players.²⁰ Figure 3(*i*) shows the development for the 5 period averages of the PEC across the third phase. The findings suggest two effects. First, the previous experience of the MFC significantly decreases the performance in the FCBC game if played under the minimal information setting. Second, the increasing payoff scheme of RI+ players significantly decreases the performance in the FCBC. We will discuss this point further below.

The effect of previous experience vanishes in the fourth phase of the experiment. More specifically, we see that players under all conditions start between 40% and 50% (i.e., RI players started with 40.5% for the

¹⁷All tests yield $p > 0.05$; Pearson's product-moment correlation test; one-sided.

¹⁸The comparison of the latter number with the frequency of the CO players is significantly smaller for $p < 0.001$ for a one-sided t-test, while all other comparisons yield $p > 0.05$ for two-sided t-tests.

¹⁹ $p < 0.001$ for one-sided t-tests.

²⁰ $p = 0.56$ for one-sided t-tests.

first ten periods, RI+ players started with 43% for the first ten periods, CO players started with 44.5% for the first ten periods, and LI players started with 50% for the first ten periods).²¹ Within the last ten periods of the fourth phase, the frequencies of PEC for RI players and for LI players increase to 64%, and 61.5%, respectively, while the frequencies for RI+ players and for CO players remain at 47.5%, and 48.5%, respectively. Note that the frequencies of the two former conditions differ significantly from those of the latter two.²² Figure 3(ii) shows the development of the 5 period averages of the PEC across the fourth phase.

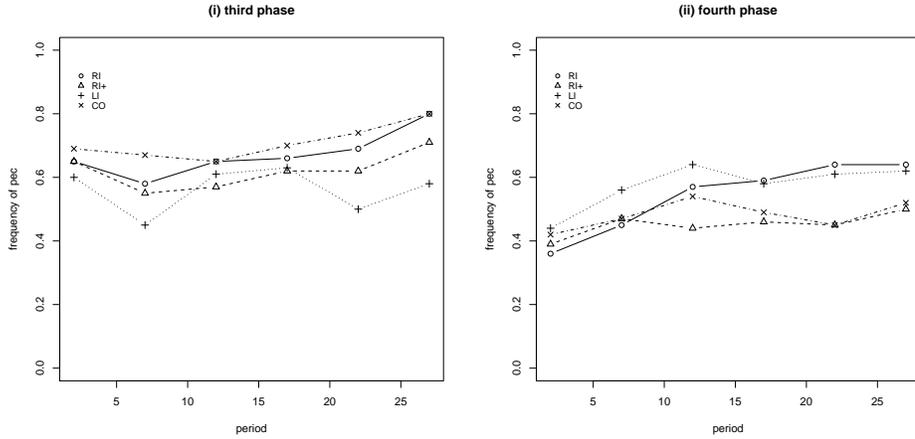


Figure 5: Average frequency of PEC choice in (i) the third phase FCBC, (ii) the fourth phase FCBC

With respect to the cumulative frequency of cells which are explored by the players in the third and fourth phase of the experiment, we find that RI players rarely explore the structure of the game. Figure 6(i) (Figure 6(ii)) shows the cumulative frequency of cells which are explored by the players in the third phase (fourth phase). Of course, this frequency is less important for LI and CO players since they cannot disentangle cells. Yet, only 30% of RI players in the third phase (35% in the fourth phase) observed all four cells of the payoff matrix, whereas 60% of RI+ players in the third phase (45% in the fourth phase) observed all four cells. Again, the results indicate that RI+ players explored the structure of the game much more accurately than RI players. However, it seems that accurate exploration does not enhance performance if played under

²¹All frequencies differ insignificantly for $p > 0.05$, two-sided t-tests.

²² $p < 0.001$, one-sided t-tests.

the rich information setting.

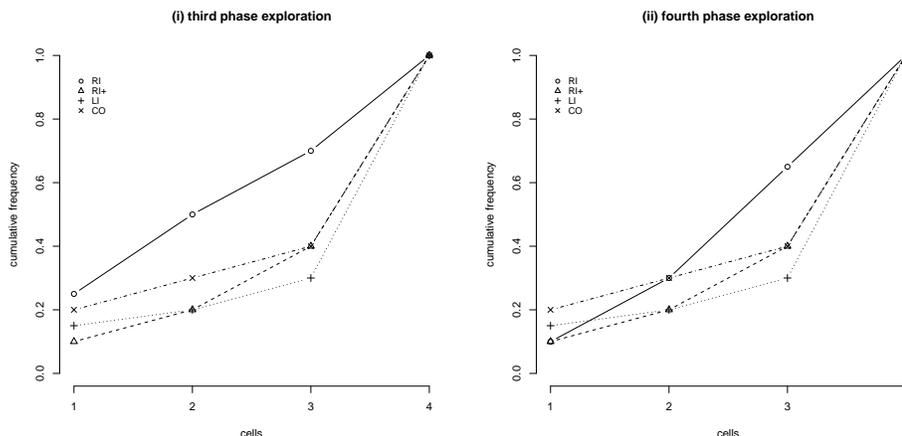


Figure 6: Cells explored in (i) the third phase FCBC, (ii) the fourth phase FCBC

We tested whether the number of sequences in which the same action has been played corresponds with the expected number of sequences under a mixture of strategies applying the Wald-Wolfowitz one-sample run tests. Yet again, we find that a significant number of players show too much inertia to support the mixed strategy hypothesis. We have to reject the mixed strategy hypothesis for 17 LI players, 28 RI players, 16 RI+ players, and 12 CO players (out of 40 players within each treatment condition) in the third phase, and for 21 LI players, 29 RI players, 12 RI+ players, and 17 CO players in the fourth phase.²³ For the analysis of the decision algorithms, we have to consider that there are two types of players in the FCBC. For one type of player, denoted as the strong player, there is a best response to the action taken by the counterpart, while for the other type of player, denoted as the weak player, any action choice is part of an equilibrium. Weak players find themselves in the same strategic situation as in the MFC since their stage game payoffs are exclusively determined by their counterparts' actions.

We run separate individual fixed effect logit regressions for both types of players, analyzing the attractiveness of action b with the latent model, as defined in equation (1). Again, choices from phase four are indicated by the dummy variable $four$. Thus, the interaction of coefficients with $four$ indicates the learning across phases. The ex ante attractiveness of b is set to 0.5. Again, we restrict our analysis to a simple autore-

²³ $p > 0.05$, Wald-Wolfowitz one-sample run tests.

	LI	RI	RI+	CO
strong players				
ρ_1	-0.048	0.812***	-0.404*	-0.89***
ρ_2	-0.399**	0.9***	0.626***	0.459***
ρ_3	0.148	4.46	-0.138	0.26
$\rho_1 \times four$	0.059	-0.342	0.505**	-0.091
$\rho_2 \times four$	0.354*	0.637***	-0.233	0.0472
$\rho_3 \times four$	-0.209	-4.255	0.453	0.174
nobs	20	20	20	20
logLik	-2134	-3687	-2160	-2194
AIC	4287	7397	4339	4408
weak players				
ρ_1	-0.238	-1.046***	-0.425**	-0.85***
ρ_2	0.45***	1.331***	0.0837	0.294**
ρ_3	0.184	-0.358	0.168	0.448**
$\rho_1 \times four$	0.456***	0.215	0.001	0.39**
$\rho_2 \times four$	0.146	0.288	0.263	0.142
$\rho_3 \times four$	-0.055	0.284	0.878***	-0.38**
nobs	20	20	20	20
logLik	-2032	-2520	-2122	-2166
AIC	4085	5060	4263	4351

Table 4: *Coefficients of the attractiveness of b in the FCBC*

gressive process with one period lag length; asterisks indicate levels of significance;²⁴ we report the number of observations (nobs); goodness of fit is reported by the log-likelihood (logLik) and the Akaike information criterion (AIC). Table 4 reports the result of our estimations.

The interpretation of the estimation results appears to be difficult. Most importantly, we find that both types of the LI players retained a myopic response in phase three, while the results for CO players show a significant ρ_1 . Only in phase four do the results for weak LI players indicate a significant ρ_1 . Thus, it seems that LI players suffer from their earlier experimental experience because they fail to structure their

²⁴* significant at a 10% level, ** significant at a 5% level, and *** significant at a 1% level.

decision algorithms in terms of the EWA. On the other hand, for RI and RI+ players we estimated decision algorithms with significant coefficients ρ_1 , both for phase three and four. Strong RI, RI+ and CO players employ decision algorithms in which ρ_1 and ρ_2 , that is, autoregressive elements and some belief learning, exercise a significant influences. Yet, while rich informed players learn across phases, that is, there are significant $\rho_2 \times four$ and $\rho_3 \times four$ for RI and RI+, strong CO players fail to do so. Rather, there is an adjustment of weak CO players to phase four, that is, $\rho_1 \times four$ and $\rho_3 \times four$. Furthermore, as in the MFC, weak RI players did not update their behavior in any significant way from phase three to phase four.

Summarizing the results for the FCBC, we find evidence that players who successfully adapted the structure of their decision algorithms to the MFC, indeed suffer from this experience in the third phase. However, the data indicate that, while the experiment continues, that is, in the fourth phase of the experiment, this previous experience facilitates PEC for players in the minimal information setting, whereas for those in a rich information setting, successful behavior does not depend on the systematic exploration of game cells.

5 Discussion

The aim of our study was to investigate whether we can find a difference in the estimated structure of a decision algorithm depending on the information received about the environment. Our results suggest that the degree of information crucially influences the way players decide in the experiment. Quite surprisingly, all decision algorithms show that, under all treatment conditions, players fall prey to an egocentric representation of the MFC. We find very simple reinforcement learning for players in the minimal information setting. Therefore, it seems plausible to assume that these players learned to respond myopically. Since we observe significant autoregressive elements in their decision algorithms, the data indicate that players in the rich information setting develop much more complex representations of the MFC. This, however, results in a curse of knowledge; more complex decision rules are less robust with respect to minor changes in the game structure, that is, if the cell permutation between the first and the second phase. As a consequence, less information facilitates Pareto-efficient coordination in the environment of the MFC, while additional information makes this more difficult. Moreover, institutional incentives, that is, the increasing payoff scheme of the plus treatment promotes the systematic exploration of the game space. However, better cognitive abilities, that is, larger short-term memory capacities, as indicated by the Wechsler digit span test, do not enhance

performance.

Yet, if the environment changes substantially, that is, the cell permutation shifts between the second and the third phase of the experiment, players in the minimal information setting suffer from the learned myopic response. Here, the more complex decision algorithm of RI and RI+ players facilitates successful coordination. Again, we find that RI players are less likely to update the structure of their decision algorithms in response to minor changes in the structure of the game, that is, across phases of the FCBC.

In summary, our data suggest that it is very difficult for all subjects to understand the games since this requires overcoming an egocentric form of representation. Rather, subjects suffer from an illusion of control. The way players decide is crucially influenced by the degree of information they receive about their environment. In general, we demonstrated that more information and a more systematic exploration of the game space lead to more complex decision rules. However, the success of simple or complex learning algorithms depends on the underlying game.

References

- [1] Camerer, C. & T.-H. Ho (1999), Experience-weighted attraction learning in normal-form games, *Econometrica*, **67**, 827-874.
- [2] Crawford, V. (1995), Adaptive dynamics in coordination games, *Econometrica* **63**, 103-144.
- [3] Conlisk, J. (1993), Adaptive tactics in games – Further solutions to the Crawford puzzle, *Journal of Economic Behavior and Organization* **22**, 51-68.
- [4] Cox, J.C., J. Shachat & M. Walker (2001), An experiment to evaluate Bayesian learning of Nash equilibrium play, *Games and Economic Behavior* **34**, 11-33.
- [5] Devetag, G. & M. Warglien (2005), Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation, *Working Paper*, University of Trento.
- [6] Erev, I. & A. Roth (1998), Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria, *American Economic Review* **88**, 848-881.
- [7] Fischbacher, U. (1999), zTree: A toolbox for readymade economic experiments, *Working Paper* **21**, University of Zurich.
- [8] Friedman, E., M. Shor, S. Shenker & B. Sopher (2004), An experiment on learning with limited information: Nonconvergence, experimentation cascades, and the advantage of being slow, *Games and Economic Behavior* **47**, 325-352.
- [9] Fudenberg, D. & D. Levine (1998), *The Theory of Learning in Games*, Cambridge, MA: MIT Press.
- [10] Greiner, B. (2004), An online recruitment system for economic experiments, in: K. Kremer & V. Macho (eds.), *Forschung und wissenschaftliches Rechnen 2003, Bericht der Gesellschaft für wissenschaftlichen Datenverarbeitung Göttingen* **63**, 79-93.
- [11] Herrnstein, R.J. (1970), On the law of effect, *Journal of the Experimental Analysis of Behavior* **13**, 243-266.
- [12] Huck, S., H.-T. Normann & J. Oechssler (1999), Learning in Cournot oligopoly – An experiment, *The Economic Journal* **109**, C80-C95.
- [13] Jordan, J.S. (1991), Bayesian learning in normal form games, *Games and Economic Behavior* **3**, 60-81.
- [14] Jordan, J.S. (1995), Bayesian learning in repeated games, *Games and Economic Behavior* **9**, 8-20.
- [15] Kelley, H.H., J.W. Thibaut, R. Radloff & D. Mundy (1962), The development of cooperation in the ‘Minimal Social Situation’, *Psychological Monographs* **76**(19).
- [16] Langer, E.J. (1975), Illusion of control, *Journal of Personality and*

- Social Psychology* **32**, 311-328.
- [17] Merlo, A. & A. Schotter (1999), A surprise-quiz view of learning in economic experiments, *Games and Economic Behavior* **28**, 25-54.
 - [18] Mitropoulos, A. (2001), Learning under minimal information: An experiment on mutual fate control, *Journal of Economic Psychology* **22**, 523-557.
 - [19] Mookherjee, D. & B. Sopher (1997), Learning behavior in an experimental matching pennies game, *Games and Economic Behavior* **7**, 62-91.
 - [20] Nagel, R. & N.J. Vriend (1999), An experimental study of adaptive behavior in an oligopolistic market game, *Journal of Evolutionary Economics* **9**, 27-65.
 - [21] Nyarko, Y. & A. Schotter (2002), An experimental study of belief learning using elicited beliefs, *Econometrica* **70**, 971-1005.
 - [22] Oechssler, J. & B. Schipper (2003), Can you guess the game you're playing? *Games and Economic Behavior* **43**, 137-152.
 - [23] Rabinowitz, L., H.H. Kelley & R.M. Rosenblatt (1966), Effects of different types of interdependence and response conditions in the minimal social situations, *Journal of Experimental Social Psychology* **2**, 169-197.
 - [24] Roth, A. & I. Erev (1995), Learning in extensive form games, *Games and Economic Behavior* **8**, 164-212.
 - [25] Schelling, T. (1960), *The strategy of conflict*, Cambridge, M.A.: Harvard University Press.
 - [26] Selten, R. & R. Stöcker (1986), End behavior in sequences of finite prisoner's dilemma supergames, *Journal of Economic Behavior and Organization* **7**, 47-70.
 - [27] Sidowski, J.B., L.B. Wyckhoff & L. Tabor (1956), The influence of reinforcement and punishment in a minimal social situation, *Journal of Abnormal and Social Psychology* **52**, 115-119.
 - [28] Thibaut, J.W. & H.H. Kelley (1959), *The social psychology in groups*, New York: Wiley.
 - [29] van Huyck, J.B., R.C. Battalio & F.W. Rankin (2005), Selection dynamics and adaptive behavior without much information, *Working Paper*, Texas A&M University.
 - [30] Wechsler, D. (1945), A standardized memory test for clinical use, *The Journal of Psychology* **19**, 87-95.

Appendix: Experimental instructions for LI players²⁵

Thank you for participating in our experiment. We kindly ask you to refrain from any public announcements and attempts to communicate directly with other participants. In case you violate this rule, we will have to exclude you from this experiment. If you do have any questions, please raise your hand and one of the persons who run the experiment will come to your place and clarify your questions.

In this experiment, you will be assigned randomly and anonymously in groups of two participants. In total, you will interact within 120 periods²⁶ with the same person. All participants received the same instructions as you did. The 120 periods²⁷ are partitioned into four²⁸ phases of 30 periods each. In each period, you (and also the other person) have to choose between two actions, a or b, and you will receive a payoff of either 0 or 1 ECU per period.²⁹ You will not be told the mechanism which determines the payoff, but we assure you that only your action and the action of the other person will determine the payoff. In particular, there is no random process controlling the payoff. The four³⁰ phases of the game are different, as the mechanism which determines the payoff for you and the other person is identical for both of you in the first phase and changes in the second, third and fourth phase.³¹ Within each phase, the mechanism remains unchanged.

After each round, you will receive the information about which payoff you earned in that round,³² as well as your accumulated payoff so far. At the end of each phase, you will be asked to describe the mechanism which controls your payoff. At the end of the experiment, we will exchange all the ECUs earned at a rate of 1 ECU = 0.08 euros, and pay off the participants.

²⁵Deviations in the other treatment conditions are indicated by footnotes.

²⁶For CO players “60 periods”.

²⁷For CO players “60 periods”.

²⁸For CO players “two”.

²⁹For RI+ players “and you will receive a payoff. The payoff in the first phase is either 0 or 1 ECU, in the second phase either 0 or 2 ECU, in the third phase either 0 or 3 ECU, and in the fourth phase either 0 or 4 ECU.”

³⁰For CO players “two”.

³¹For CO players “and changes in the second phase.”

³²For RI and RI+ players additionally “the action the other participant chose.”.