Learning under minimal information: an experiment on mutual fate control

Atanasios Mitropoulos

Faculty of Economics and Management
Otto-von-Guericke-University Magdeburg
PO-box 4120, 39016 Magdeburg, Germany
e-mail: atanasios.mitropoulos@ww.uni-magdeburg.de

April, 2000

Abstract:

Reinforcement learning has proved quite successful in predicting subjects' adjustment behaviour in repeatedly played simple games. However, reinforcement learning does not predict convergence to the efficient cell in the minimal information game of mutual fate control, while earlier psychologists' experiments show some tendency to convergence. Our rivalling learning rule, a modification of win-stay lose-change, does predict convergence. We perform an experiment using modern economic methodology and compare these two learning rules. Our results are unfavourable for both reinforcement learning as well as win-stay lose-change. The data rather support the view that subjects search by using patterns.

Keywords: mutual fate control, learning, coordination, experimental economics, coordination failure

JEL classification: C72, C92

I am grateful to Jeannette Brosig and Joachim Weimann for valuable comments.

1. Introduction

During the last decades literature on learning in games has been growing fast. While in the early stages of research on learning focus has been laid on the justification of equilibria (see Brown 1951, Robinson 1951), during the last years learning rules have increasingly been applied to predict individual behaviour (see Crawford 1995, van Huyck et al. 1996, Mookherjee and Sopher 1997, Erev and Rapoport 1998, Tang 1998, and many more). Notably, learning theories of limited cognitive sophistication have been quite successful in predicting individual behaviour (see especially the reinforcement model of Roth and Erev 1995, and Erev and Roth 1998)[1], even though players in experiments were given much more information than actually being used by these theories. Alternative learning theories that make use of more information, such as fictitious play[2], have regularly been outperformed (see for example Roth and Erev 1995, Mookherjee and Sopher 1997, Camerer and Ho 1999). These results confirmed Smith's (1990) early conjecture that individuals rigorously select among pieces of information given to them and specified the most relevant piece of feedback to be the own payoff[3].

There are only few experimental studies that deal with environments in which the own payoff from the last interaction is the only feedback individuals get. The results from these studies are mixed. Mookherjee and Sopher (1994), for example, notice that in a game involving a unique mixed-strategy equilibrium subjects are showing much more inertia when feedback about the opponent's performance is lacking. Nagel and Vriend (1997) show that in a fairly complex situation subjects adhere to a simple directional learning scheme. Van Huyck et al. (1996) find that in a coordination game with a large strategy space fictitious play does a better job in predicting the speed of convergence than does reinforcement learning, although subjects did not have enough information to perform fictitious play.

Our purpose is to investigate, whether reinforcement learning is able to make good predictions of individual behaviour in a most favourable setting. We take one of the simplest possible social games and impose a minimal informational setting which allows subjects to act according to this learning scheme. In section 2 we present the game and discuss previous results from psychological experiments. Also in section 2 we present results from simulations

[1] However, the opposite result has been found in van Huyck et al. (1996) and Cheung and Friedman (1997).
[2] Fictitious play has originally been introduced by Brown (1951) and Robinson (1951). It has extensively been discussed in Fudenberg and Levine (1998).
[3] There are quite some studies, however, that also show that feedback on actions or payoffs of the opponents have significant impact on one's own actions, too. See for example, Huck et al. (1998, 1999) and Duffy and

that indicate that reinforcement learning seems to predict different behaviour than observed in the earlier studies. Therefore, we also present an alternative learning sch eme that seems to perform better in accounting for the results.

Section 3 is devoted to the experiment design, while section 4 presents and discusses the data. We will see that behaviour in our experiment differs from what has been observed in psychological studies. Furthermore, none of the learning schemes successfully describes all broad qualitative features of actual behaviour. The general result is that subject's play is much more complex than can reasonably be traced by simple adaptive learning schemes. For this reason the game of mutual fate control provides an environment in which simple dynamics would lead to efficiency, but subjects are handicapped by the complexity of their own thoughts. We discuss our results and conclude in section 5. Appendix 7.1 contains the instructions given to the subjects (translated from German), and appendix 7.2 shows all observations. Finally, appendix 7.3 gives detailed information on the accordance to the experimentation learning scheme.

2. Game and predictions

In this section we will present the game that has been termed *mutual fate control* (Thibaut and Kelley (1959)). The first subsection will discuss its general characteristics. In the next two subsections we will show that reinforcement learning and experimentation learning make two distinct predictions for long-run play.

We tried to find a game from which we can expect learning theories to be particularly easy to distinguish. We considered all 2x2-games that involve only two possible payoff values, i.e. either 0 or 1. The most interesting game that we were left with and which suffices to make distinct predictions for different learning schemes was the game known as mutual fate control. Psychologists are more familiar with this game than are economists, since it served as the basic game representing a *minimal social situation*, i.e. one of the simplest situations in which at least two persons are involved and in which at least one strategy of one player affects the payoff of at least one other player, possibly without being aware of this. However, even psychological experiments date back as far as 1981 with the peak of interest having been during the early 60s. This is why we decided, first, to replicate the experiment in modern style, using current economic experimental methodology, and second, to describe the game in more detail in the next subsection. We also give a brief survey of the psychological results.

Feltovich (1997). However, there have opposing results been presented, too (Bosch-Domènech and Vriend 1999).

2.1 The game

The game under consideration is an extremely simple 2x2 game with payoffs being either 0 or 1. The essence of the game is that each player's payoff is solely determined by the action taken by the opponent. In our simple 2-value 2x2 version this means that each player decides upon whether to give 0 or 1 to the other player. The resulting payoff bimatrix is shown in Table 1.

		player 2	
		A	B
player 1	A	0 0	1 0
	B	0 1	1 1

Table 1: mutual fate control

An extension to continuous strategy spaces is straightforward. For three reasons we deliberately chose the simplest version of this game. First, we wanted to avoid possible sampling processes due to the complexity of the strategy space. Second, in a more simple game it is most likely that the underlying structure of the game is quickly revealed by repeated play. And third, by this design we can rule out endogenous aspiration levels that may arise from the exploration of a larger strategy space[4]. As can easily be seen, under complete information any mixed-strategy profile is a weak Nash-equilibrium. People are, hence, faced with an extreme multitude of equilibria. However, since there is only one efficient cell (which also gives the same payoff to both players) any motivational theory will suggest that individuals will prefer this cell (B,B)[5].

With uncertainty about the payoffs, however, the underlying payoff structure has first to be investigated in order to be able to coordinate on cell (B,B). Since no information of either actions or payoffs of the opponent is given to the players, coordination has to emerge *silently*, i.e. individual calculus does not suffice for coordination, while joint action can lead to efficiency. In the next two subsections we show that both, divergence as well as silent coordination, is possible under different individualistic learning rules.

In earlier experimental investigations, psychologists have found strong evidence in favour of successful cooperation (Sidowski et al. 1956, Sidowski 1957, Kelley et al. 1962,

[4] Inclusion of endogenous aspiration levels may cause learning to have distinctly different characteristics. See Börger and Sarin (1997).

[5] We consider it as highly improbable that people exhibit so much spite as to deliberately harm the opponent and to run the risk of immediate retaliation, thereby giving up the jointly and individually efficient outcome.

Rabinowitz et al. 1966, and Arickx and Van Avermaet 1981). The authors report that usually, coordination after 100 rounds of repeated play was reached to approximately 75 per cent. However, they usually did not tell the subjects that they were actually playing a 2-person game[6], and second, incentives were generally non-monetary; instead, electric shocks provided negative payoffs while points (not converted into money) provided positive incentives.

In more detail, Sidowski et al. (1956) first conducted mutual fate control under a free-timing scheme, i.e. subjects could choose among strategies at any time they wanted. The result of each choice was immediately forwarded to the opponent. The common observation was that cooperative play rose gradually over time, despite of players not being aware of participating in a social environment. Sidowski (1957) replicated this finding.

Experiments by Kelley et al. (1962) showed that the timing scheme seems to be crucial. They found that mutual fate control leads to successful cooperation if responses are made simultaneously, but cooperation does not evolve if responses have to be made in alternating order, i.e. the outcome of one player's response is showed to the other player before he is allowed to do his own choice. Kelley et al. argue that cooperation is the result of a win-stay lose-change learning scheme, which is successful in a simultaneous-move environment (or an environment in which simultaneous moves are frequent), but yields misleading incentives in an alternating-move environment. Furthermore, from individual behaviour Kelley et al. found that the win-stay part of the learning scheme is strong while the lose-change part cannot be confirmed. Rabinowitz et al. (1966) report confirming results from similar experiments on mutual fate control and a variant, termed *fate control – behaviour control*, by using a framing in which subjects' choices were presented as predictions of a bivariate random variable and the other players' responses were presented as successes or failures.

From the point of view of today's accepted methodology all experiments suffer from severe shortcomings. First, the early ones gave incentives to opponents to maximise one's own positive payoffs (points), while negative payoffs (electric shocks) were seen as negative stimuli per se. Given the fixed time limit (usually five minutes) and no restrictions on the speed of responses, people tended to hit the buttons as often as possible, yielding response frequencies of several hundred times within the time interval of up to five minutes. Second, response frequencies were recorded "on-the-fly", i.e. every 30 seconds the experimenter had to take down four numbers, representing the four responses of the two subjects, simultaneously. At the given response speed this task was certainly lacking accuracy. Third, because of the free timing scheme it is not credible that subjects were not aware of interacting

[6] So subjects had the idea of playing against a two-armed bandit.

in a social environment. People must have quickly found out that no random variable was determining their responses but a human person, even more so because, at that time, electronic techniques were not very much advanced.

The problem of making subjects believe that they do not interact with other subjects was not solved by introducing simultaneous or alternating moves, because there is no reason to control for the timing of responses other than coordinating responses with responses of other subjects[7]. However, the controlled schemes allowed for more detailed individual analyses and are much closer to today's experimental practice in economics. We basically build on the simultaneous-move experiment by Rabinowitz et al. (1966) with some modifications that make our experiment adhere to current experimental designs in economics. Especially, we do not try to conceal the social interaction. Instead, we tell subjects in advance that they are going to play a two-person game. Furthermore, we give true monetary incentives by paying subjects according to their payoffs gathered during play. Finally, we took great care to make interaction anonymous, so persons, first, cannot make use of known personal characteristics of their opponent, second, communication during play is not possible, and third, personal attitudes towards their opponent do not play a role.

2.2 Reinforcement learning

The results from former experiments on the simultaneous-move mutual fate control game are quite favourable for the prediction that eventual cooperation is quite probable. We now check whether reinforcement learning, the most prominent and most successful learning scheme that is applicable to the given low-information environment yields the same predictions. We describe the reinforcement model as has been introduced by Roth and Erev (1995)[8].

The starting point is to think of probabilities of choices being the normalised inclination to choose a certain strategy. This tendency to play a strategy is determined by the accumulation of past payoffs from choosing that strategy. In formal terms we say that $p_j^i(t)$ is the propensity of player *i* to play strategy *j* at time *t*. The initial propensity $p_j^i(0)$ has to be determined either by fixing it in advance, or by estimating the value. Let $p^i(t)$ be the payoff

[7] The only experiment to fulfil the assumption of a minimal social situation were the experiments by Arickx and van Avermaet (1981). They replicated the experiment by Rabinowitz et al. by using a programmed opponent who responded according to the Cross (1973) dynamic.

[8] Psychologists have discussed reinforcement learning long before economists have taken it up for their own purposes. The first description in economic terms has been made by Cross (1973), though his model was slightly different from the one by Roth and Erev (1995).

to player i at period t and $a^i(t)$ be the action chosen. Then the updating of the propensities follows

$$p_j^i(t+1) = \begin{cases} j \cdot p_j^i(t) + p^i(t) & \text{if } a^i(t) = j \\ j \cdot p_j^i(t) & \text{otherwise} \end{cases}, \tag{2}$$

with $j > 0$ being the forgetting parameter[9]. The probability of strategy j being chosen by player i at time t is then given by

$$s^i_j(t) = \frac{p_j^i(t)}{\sum_{j=1}^{J^i} p_j^i(t)}. \tag{3}$$

Denote player i's probability distribution over his set of strategies by $s^i(t) = \left(s^i_1(t), \ldots, s^i_{J^i}(t)\right)$. Then the action by player i at period t is determined by

$$a^i(t) = Z\left(s^i(t)\right)$$

where $Z(s)$ is a random variable with distribution s.

In order to present the characteristics of reinforcement learning in our particular game we proceed in the following way. First, in order to get an idea of how reinforcement learning behaves in the long-run we present results from simulations after 1000 rounds of play. We then compare these results with simulations of 100 rounds. We will see that results from the two time-horizons differ only with respect to a certain range of the value for the forgetting parameter j.

One of the most prominent features of reinforcement learning is the adherence to the *law of effect*. Roughly, this effect states that learning is fast at the beginning of play and gradually diminishes as time proceeds. For our 2x2 game this means that after an initial phase of rapid adjustment subjects will slowly settle to some mixed-strategy profile. As we see from simulations this process results in extremely diverse strategy profiles covering almost all of the set of mixed strategies. Figure 1 depicts 10000 profiles resulting from 1000 periods of repeated play by reinforcement learners. We set initial propensities equal to 1 for each strategy and each player and j = 1. Each dot of the figure represents the probability of either player to play strategy A in the 1001st period.

[9] See Roth end Erev (1995) for a discussion of this parameter.

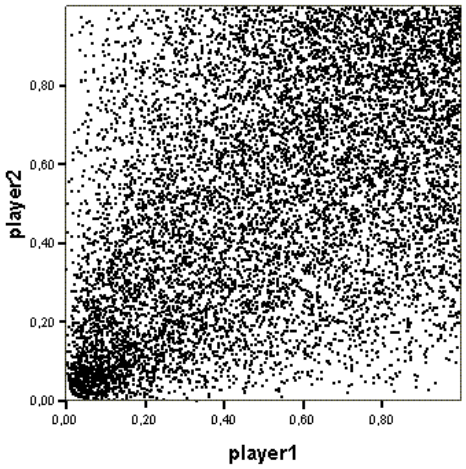


Figure 1: convergence of profiles (s_A^1, s_A^2) with reinforcement learners

We see that our virtual subjects have a slight tendency to either move to the efficient state (0,0) or to remain near the Pareto-worst state (1,1). However, even after 1000 rounds behaviour is extremely diverse.

Things change when considering forgetting, i.e. when $j < 1$. We summarise our findings from simulations of long-run behaviour (i.e. the resulting choice probabilities after 1000 rounds) in Table 2. For convenience we classify choice probabilities in those with $s_A \approx 0$, $s_A \approx 0.5$, and $s_A \approx 1$, where s_A denotes the probability to play strategy A. What cannot be seen from the table is that for any value $j < 1$ the classifications are very precise, i.e. the virtual subjects eventually converge to either of the classes. For $j = 1$ we get the diverse picture as illustrated in Figure 1[10].

profile	$j = 0.1$	$j = 0.5$	$j = 0.9$	$j = 0.95$	$j = 0.99$	$j = 1$
(0,0)	3.6	17.0	32.6	32.1	25.0	22.7
(0,0.5)	58.9	35.7	4.7	3.8	7.7	16.5
(0.5,0.5)	0.0	0.0	0.0	0.0	7.7	12.7
(0,1)	3.7	15.3	29.3	30.3	17.3	8.4
(0.5,1)	0.0	0.0	0.0	3.0	24.1	23.4
(1,1)	33.8	32.0	33.4	30.8	18.2	16.3

Table 2: long-run profile classification (in per cent) for reinforcement learning; 1000 simulations with 1000 rounds each

[10] For the case $j = 1$ it is important to know that the boundaries of the classes were defined at 0, 0.33, 0.67, and 1.

From Table 2 we can infer that probability distributions of states show highly non-linear dependence on the forgetting parameter j. The most important characteristics to be observed are (i) that the probability of the efficient state (0,0) to be approached never exceeds 33 percent, and (ii) that profiles in which at least one player mixes with 50 percent probability look rather unstable, so they can survive after 1000 rounds only if forgetting is small enough.

Simulations on the period of time that was actually being played in the laboratory (i.e. 100 rounds) is shown in Table 3. The basic result from this analysis is that after 100 rounds things are not much different from what they are in the long run, unless the forgetting parameter j is not close to 1. However, from earlier analyses we may expect j to lie within this region (see for example Camerer and Ho 1999). In this case we can see that in the shorter term strategies with mixing behaviour of at least one of the participants are more likely to occur than in the long-run.

profile	$j = 0.1$	$j = 0.5$	$j = 0.9$	$j = 0.95$	$j = 0.99$	$j = 1$
(0,0)	3.8	19.0	31.6	27.7	20.9	18.2
(0,0.5)	60.4	35.0	8.4	7.5	15.8	19.0
(0.5,0.5)	0.0	0.0	0.2	2.7	13.7	14.2
(0,1)	3.2	15.9	19.5	15.9	9.7	8.2
(0.5,1)	0.0	0.0	9.6	22.3	24.1	25.3
(1,1)	32.6	30.1	30.7	23.9	15.8	15.1

Table 3: short-run profile classification (in per cent) for reinforcement learning; 1000 simulations with 100 rounds each

2.3 Experimentation learning

Since reinforcement learning is not quite satisfying as to the behaviour that can be expected from the results of earlier studies we next present a learning rule that is purely based on experimentation behaviour. We are unaware of any former formulation of this rule, though it bares resemblance to the search-and-select adaptation rule proposed by Conlisk (1993). It may also be interpreted as a variant of the win-stay lose-change learning scheme, whereby the win-stay part is slightly generalised and the lose-change part is transformed into a lose-randomise scheme.

The idea of this rule is to let individuals, first, gather information about payoff and then decide upon which action to take, based on a fixed number of previous payoff experiences with each strategy. In more detail, we assume each player to record the feedback returned after playing a strategy and keep only the last m pieces of feedback information for each strategy. In case the record for one strategy is unambiguously favourable then this strategy

will be chosen with a fixed probability x. In case no action or both actions have a perfect record, the action chosen in the next period is determined according to a random process giving equal probability to each possible strategy.

There are several different ways to assess the goodness of a strategy record. Since we want to apply our learning rule to a particularly simple game we do not bother to formulate a complicated evaluation scheme. In the game to be described, the only payoffs that possibly arise from choosing an action are either 0 or 1 (and people are informed of that prior to playing the game). This simplicity allows us to formulate a rudimentary evaluation scheme based on a fixed aspiration level. The aspiration level is given by the payoff 1. Hence, if the payoff of the action is 0 we record dissatisfaction (or failure), and if the payoff of an action turns out to be 1 we record satisfaction (or success).

We will concentrate on a 2-person game. Players are denoted by $i \in \{1,2\}$. The number $m \in N$ denotes the size of the memory of a player. For the later estimation we restrict m to be equal for all players, though for the theory we might as well allow it to vary between individuals. Our notion of the size of the memory m is used differently to usual definitions of bounded history, as for example in Hon-Snir et al. (1998) and Milgrom and Roberts (1990, 1991). While bounded history is defined as capturing all payoff feedback dating back m periods before the current period, we consider the player to trace all m last actions *of each strategy* separately. Particularly, if a strategy has not been chosen for a long period of time, then even experience gathered in the distant past is relevant for the current choice. This model is related to the view that memory is not simply forgotten, but rather selectively overwritten by new experience. The stack of memory of player i is given by the vector $s^i = \left(s^i_1, \ldots, s^i_{j^i}\right)$ of strategy records s^i_j. A strategy record at time t for strategy j of player i is the vector $s^i_j(t) = \left(e^i_j\left(t^i_{j,t-m+1}\right), \ldots, e^i_j\left(t^i_{j,t}\right)\right)$ of evaluations $e^i_j \in \{0,1\}$ at periods $t^i_{j,t}$ in which strategy j has been chosen by player i for the t^{th} time ($t \geq$ m). The value of the evaluation $e^i_j(t)$ is given by 1 if the payoff for player i resulting from playing strategy j in period t has been equal to 1[11], otherwise it is 0.

The experimentation learning is now defined to accord to the following rule:

---

[11] More generally, we may define the value of $e^i_j(t)$ as a success (i.e. equal to 1) whenever the resulting payoff from playing strategy j has been bigger or equal to player i's aspiration level a^i, and we fix $a^i = 1$ for all players i. For the present study it is not necessary to further model aspiration levels. A generalisation for more complex payoff spaces as has been done by Börgers and Sarin (1997) for the Cross model can easily be applied.

$$a^i(t) = \begin{cases} Z\left(x \cdot 1_j + \frac{1-x}{J^i - 1} \cdot 1_{-j}\right) & \text{if } \left[s_j^i(t) = 1^m\right] \wedge \left[\forall g \neq j : \left(s_g^i(t) \neq 1^m\right)\right] \\ Z\left(\frac{1}{J^i} \cdot 1\right) & \text{otherwise} \end{cases} \quad (1)$$

whereby 1_j denotes the vector with all entries being 0 except at the j [th] entry which is 1 and 1_{-j} being the complementary vector. 1^m denotes the *m*-tuple consisting entirely of ones, and $Z(s)$ again being the random variable with distribution s.

A very interesting extreme case is x = 1, i.e. experimentation with equal probability until one strategy gets a full record. The crucial property of this learning process is that play among two players following this process eventually converges to the efficient state. The reason for this is simply that random play by both players will never cause both players to commit to some pure strategy simultaneously, unless both players have a record with complete successes. Under the 2x2 mutual fate control game this is only possible if both players have reached the state at which they are playing strategy B with certainty. Furthermore, since probability distributions under non-commitment are assumed to stay fixed over time and to have full support, the outcome (B,B) will eventually occur with probability 1 [12].

For illustration we ran simulations with x = 1 in order to assess the distribution of the probability of convergence over time. Probabilities of convergence strongly depend on the size *m* of memory. Figure 2 shows the cumulated distributions of date of convergence. They are composed of 10000 simulations for each $m \in \{1,2,3,4\}$.

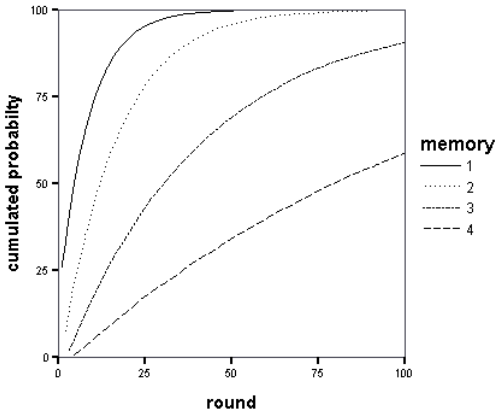


Figure 2: Distribution of convergence for experimentation learning

Not surprisingly, convergence quickly shifts to the later rounds as m increases. We will have to bear this in mind when interpreting our data. However, for $m \leq 3$ convergence until period 100 is quite likely.

[12] From the point of view of Markov Chains the argument is that the state, at which both players have a full success record for strategy B, is the only absorbing state of the game.

3. Experimental procedure

The experiment was conducted in four sessions at the MaxLab[13]. In each session 8 to 10 subjects, most of them business and economics students at various levels, were randomly allocated to computer terminals, which were separated by mobile cardboard devices. By internal computer assignment subjects were then randomly matched to pairs. A custom made computer program, written in Java, helped to make decisions conveniently and to view all information gathered during play. There were 17 pairs (34 subjects) participating in the experiments. Below we describe which information was given.

Each pair played 100 repetitions of the stage game shown in Table 1. Players did not receive any fixed amount of money for participation. Instead, their entire payoff was determined during the game. A payoff of 0 lab dollars was converted into 0 Deutsche Mark and a payoff of 1 lab dollar was converted into 0.30 Deutsche Mark. Participants were paid at the end of the experimental session, whereby fractions of Deutsche Marks were rounded to the next higher integer. So minimum and maximum payoffs that could be earned were 0 Deutsche Mark and 30 Deutsche Mark, respectively. On average subjects earned 18.25 Deutsche Mark, whilst a session lasted for 45 to 60 minutes. However, large differences between individual total payoffs indicate that considerable payoff incentives were given. The maximum total payoff was 30 Deutsche Mark, while the minimum total payoff was only 2 Deutsche Mark.

In contrast to many psychological experiments we followed the economic convention to inform subjects, before starting to play, about them being involved in a two-person game. We did this mainly in order to avoid uncontrolled beliefs on part of the subjects about the interaction they are going to participate in. Furthermore, this keeps consistency with almost all economic experiments. After that, subjects were also told that each participant got the same instructions. Furthermore, subjects knew about playing a game in a pair with an opponent staying fixed for the entire duration of the experiment. They knew that they would never get any information about their opponent, neither his payoffs, nor his choices, nor his identity.

The information about the payoff structure of the game was minimal. Subjects were left ignorant about the entries of the payoff matrix. However, the instructions clearly indicated that payoffs were calculated with reference to a payoff table that remained fixed over time.

[13] The Magdeburg Experimental Laboratory, Germany.

Furthermore, it was common knowledge that payoffs for any strategy combination of a pair could take on either of the two values, 0 or 1[14].

4. Data analysis

4.1 Summary statistics

Average choice frequencies, at first sight, seem to be in line with the former findings of the psychological literature. As Figure 3 shows, the 10-period moving average of the rate of B-moves shows a slight upward trend over time, while choices obviously contain a large amount of variance. This is confirmed by a simple linear regression of mean choices on time, where the coefficient for the time variable is positive and significant to the 1% -level. The rate starts at roundabout 54% and approximates roughly 66% towards the end. Final coordination is not as high as in Kelley et al. (1962), but it is considerably more than the benchmark case of complete randomisation, i.e. 50%. So far, it looks like slow adjustment process and gradual awareness of the coordination scheme are confirmed by the data. Below, we will see that both conjectures are falsified by the structure of individual play.

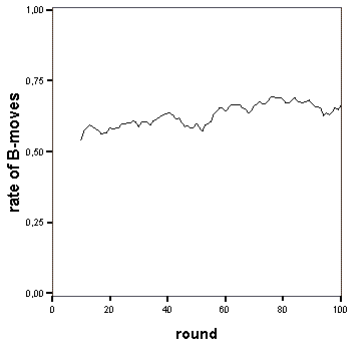


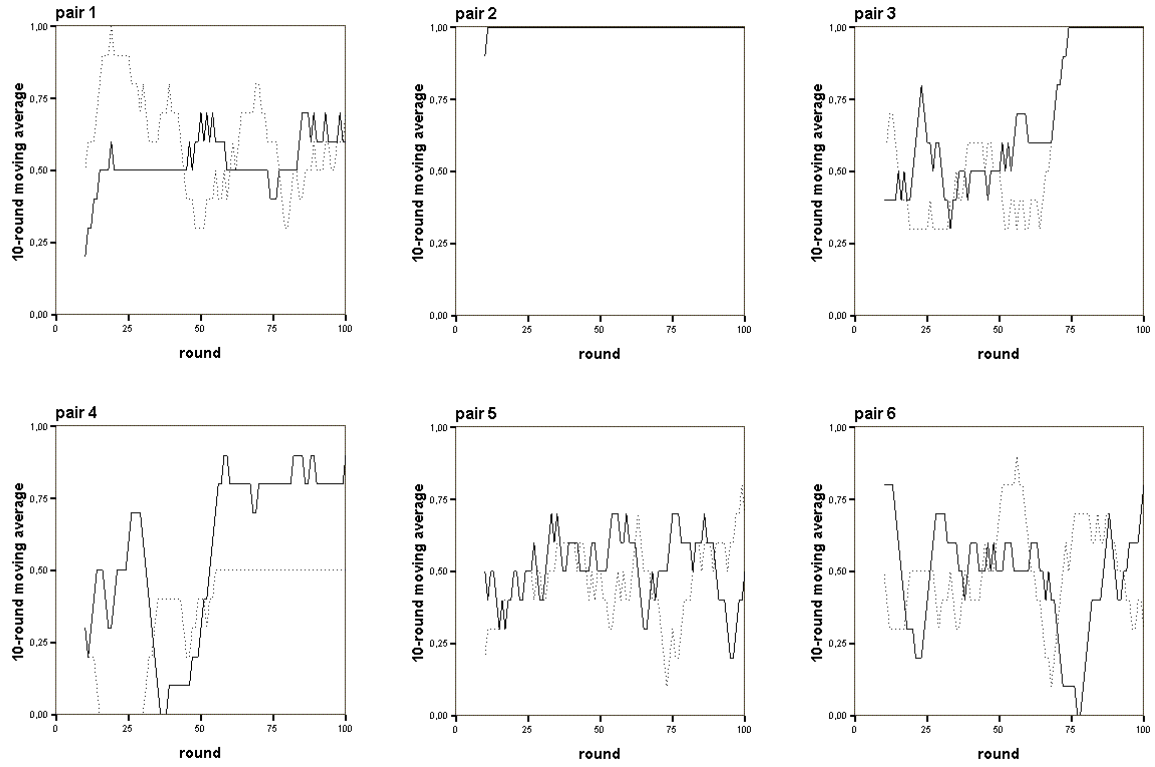
Figure 3: 10-period moving average of rate of B-moves (all players)

We start our more detailed analysis by having a look at the number of pairs that show successful coordination: seven pairs (pairs 2, 3, 9, 10, 13, 16, and 17) eventually coordinate to the efficient cell while ten pairs fail to do so. This is a clear rejection of the hypothesis that people use experimentation learning with high x and low memory (i.e. $m \leq 3$)[15]. The rate of successful coordination is simply too low. Furthermore, a similar rate of convergence is what has been expected by reinforcement learning with slight forgetting (i.e. $j = 0.9$).

[14] Instructions are given in appendix 7.1.

[15] In their corresponding treatment Rabinowitz et al. (1966) find less complete coordination. The number of coordinating pairs was 3 out of 19. However, in a slightly different experiment Kelley et al. (1962) had coordination rates similar to ours, namely 11 out of 30 pairs in a first experiment and 12 out of 22 pairs in a second experiment.

A closer look at those pairs, which eventually coordinate on playing the efficient cell, again draws a different picture. Five of the seven pairs agree on playing the efficient profile (B,B) no later than the ninth round (pairs 2, 9, 10, 13, and 17). One pair starts playing (B,B) as late as round 65 (pair 3), while only one pair shows a gradual adjustment towards the efficient cell. From this, we may infer that a slow adjustment process, as is predicted by reinforcement learning, is not typical of actual play. Having another look at the one pair that converges rather late in the game we find that a subtle harmony in choices, assum edly born out of pure chance, preceded convergence. The sequence of actions from round 58 to 65 was (B,A), (A,A), (A,A), (B,B), (A,A), (B,B), (A,A), (B,B), on which (B,B) followed thereafter. And even in pair 16 which does not show any explicit point of co nvergence B-play seems to have come up rather suddenly. This can be seen from Figure 4 which shows the empirical frequencies of choices as a moving average of 10 rounds for each player. Hence, in case pairs do converge, we find hardly any evidence for slow adjustment behaviour as is predicted by reinforcement learning.



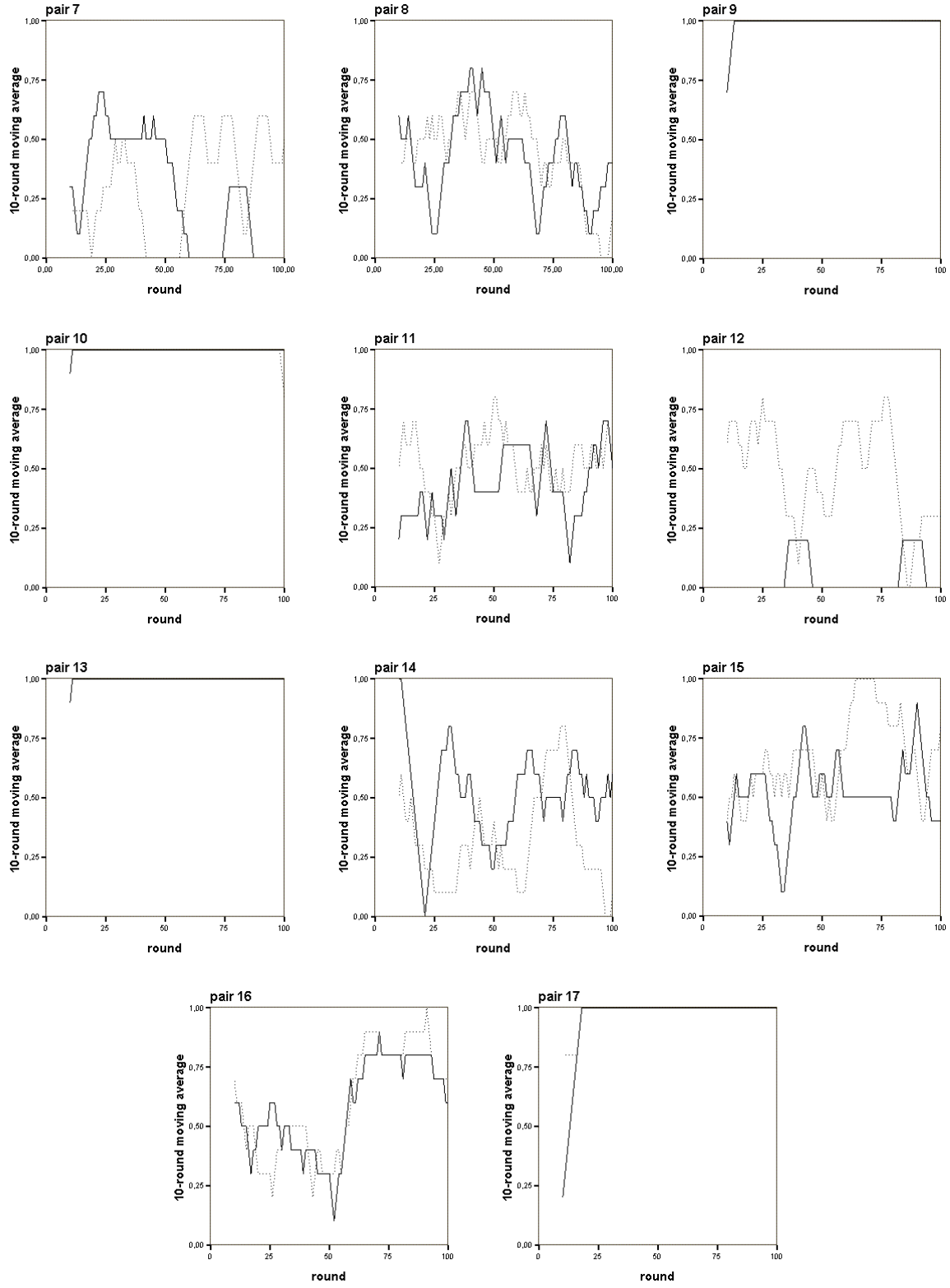
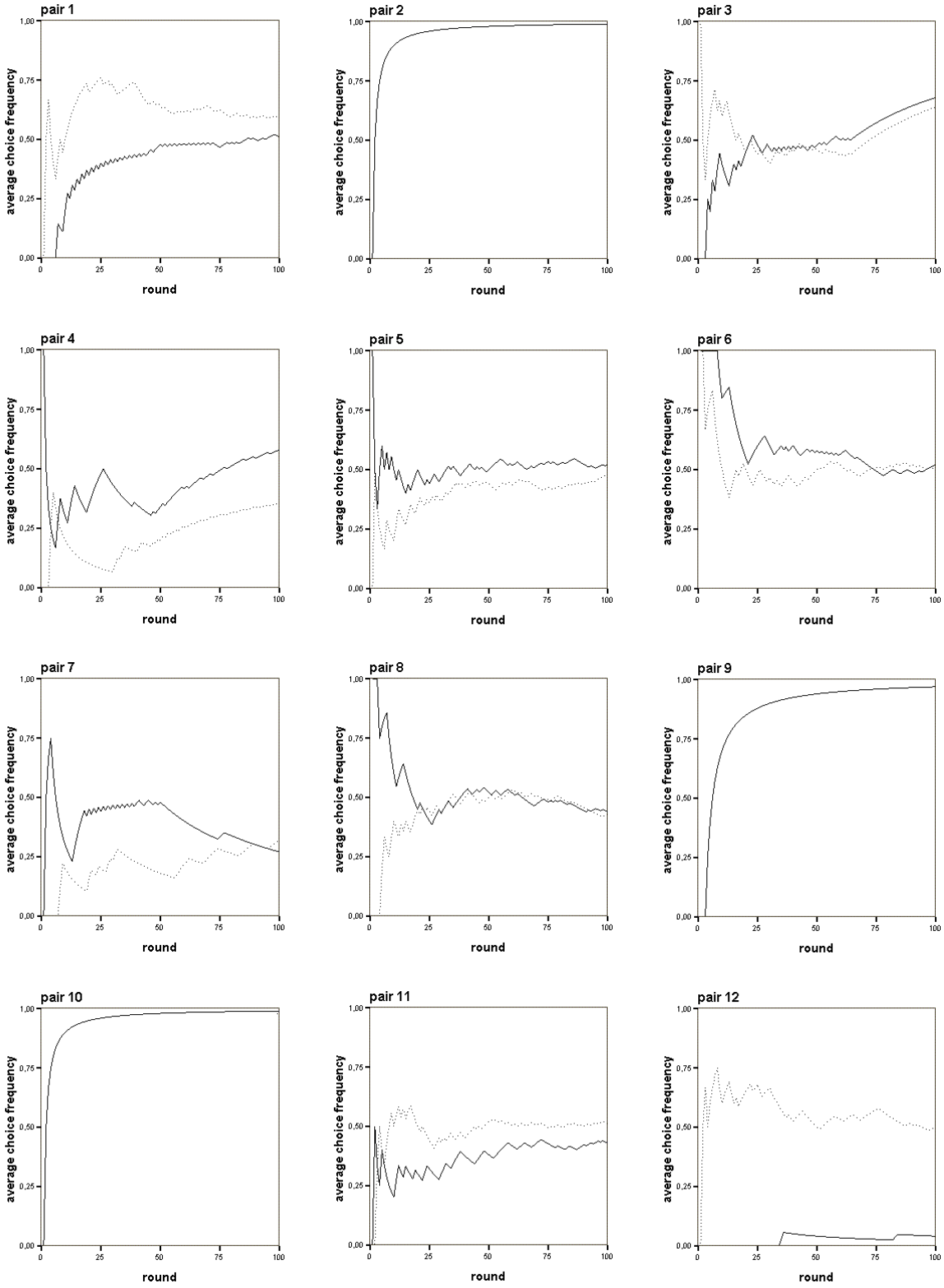


Figure 4: individual 10-round moving averages of B-frequencies

In fact, slow adjustment can neither be found in those pai rs that did not converge. As can be seen from Figure 4, behaviour is rather volatile over time, within subjects and across subjects. Furthermore, as can be seen from Figure 5, for eight out of the ten pairs that do not converge to the efficient outcome, empirical frequencies over all periods show a tendency

towards equal frequency for either strategy. This is contrary to what we observed in our simulations on reinforcement learning in section 2.2.



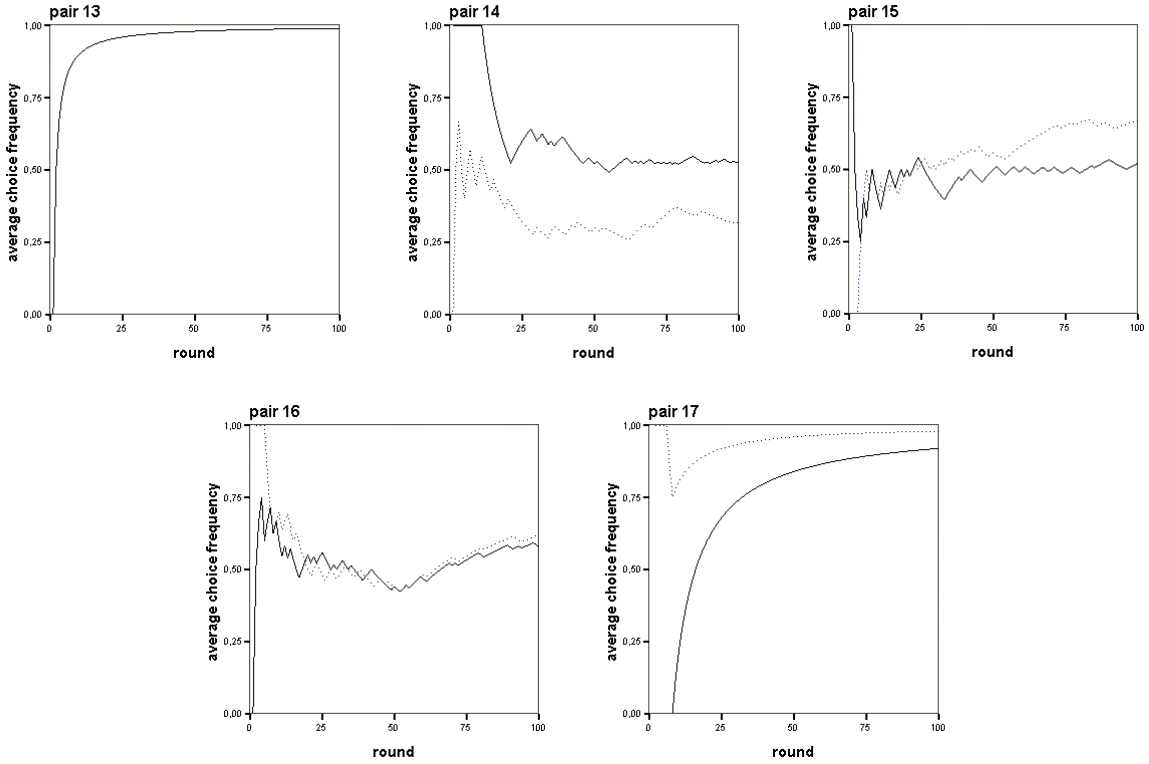


Figure 5: individual average choice frequencies over time

We can see this also by taking Figure 3 and removing the decisions made by those fourteen players who eventually coordinate with their partners, leaving us with the 10 -period moving average of the rate of B-moves over time only for those players who do not eventually coordinate with their partners. This is shown in Figure 6. The same simple linear regression as before now tells us that the coefficient for the period variable is positive but so small (as compared to its variance) that significance cannot be established even to the 10% -level.

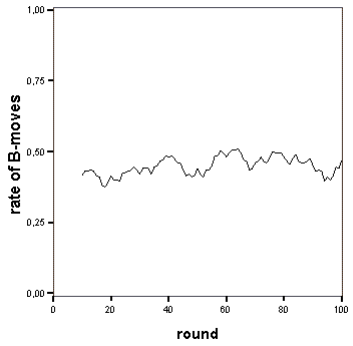


Figure 6: 10-period moving average of rate of B-moves
(only players in non co-ordinating pairs)

Finally, we checked for the accordance to the experimentation learning scheme. For each memory size *m* we computed the rates at which strategies were kept after having a full rec ord

of confirmations and the rates at which strategies were changed when the record was ambiguous, i.e. contained at least one failure. These rates for each strategy as well as for both strategies are depicted in Table 4. According to the experimentation learning scheme with x = 1, we expect staying when the record is full of confirmations to occur with probability 1 and changing when the record is ambiguous to occur with probability 0.5. Note that for $m = 1$ we can directly compare behaviour with the win-stay lose-change strategy which has, as yet, been proposed to account best for behaviour in the mutual fate control game (see Kelley et al. (1962) and Rabinowitz et al. (1966))[16]. As in the experimental learning scheme, the win-stay lose-change scheme would predict a win-stay rate of 1. However, the lose-change rate would also be predicted to be 1.

Dynamic	response	m=1	m=2	m=3	m=4
full record-stay rate	both	0.63	0.63	0.57	0.54
ambiguous record-change rate	both	0.42	0.40	0.40	0.40
full record-stay rate	A	0.65	0.65	0.60	0.62
full record-stay rate	B	0.47	0.52	0.44	0.29
ambiguous record-change rate	A	0.42	0.41	0.40	0.40
ambiguous record-change rate	B	0.45	0.44	0.44	0.44

Table 4: accordance with response dynamics of experimentation learning (averages over all subjects in non coordinating pairs)

From the figures for both responses taken together we clearly see that players are far from according to experimentation learning. Best approximation is reached for $m = 1$ and $m = 2$, while higher values for m seem to predict even worse. In line with earlier observations, we find that staying with a strategy after it has been confirmed is much more likely than changing one's strategy after failure. However, contrary to former studies, changing after failure occurs with considerably lower frequency than 0.5, which is a complete rejection of the lose-change prediction and is even significantly lower than predicted by random play [17]. Furthermore, for $m = 1$ and $m = 2$, the sum of the first two rows only marginally exceeds 1 which indicates that staying and changing behaviour is almost independent of whether the strategy's record is confirmatory or not. We conclude that our subjects show a significant amount of inertia, but the amount of inertia is almost independent of whether strategies have a confirmatory record or not.

[16] Note that in the mutual fate control game Selten's directional learning theory reduces to the win-stay lose-change strategy.

[17] The detailed analysis runs as follows: for each player participating in a non-converging pair we use the binomial test to determine whether the loose-change decision is (to the 5%-level) significantly lower or higher than 0.5. For those players showing significance we find that 7 players change less often than 0.5 and 2 players change more often than 0.5. According to the binomial test this distribution is (to the 5%-level) significantly biased towards lower frequencies than 0.5.

Looking at the data for each strategy separately, we see that, ironically, a confirmatory record for strategy A causes subjects to stay with larger probability than a confirmatory record does for strategy B. We have no account for this observation. The sudden drop of the rate of staying when the record is full at strategy B is due to few observations [18].

The discouraging results from Table 4 stand in sharp contrast to results from earlier studies in which the win-stay lose-change strategy has been put forward to account for observed behaviour. Rabinowitz et al. (1966) showed that in their experiment as time proceeds subjects increasingly accord to the win-stay lose-change strategy. We recalculated their table 2 with our own data and found no such adjustment behaviour [19].

		rounds	
tendency	response	2 - 50	51 – 100
win-stay	both	0.62	0.57
lose-change	both	0.42	0.38
win-stay	A	0.68	0.63
win-stay	B	0.55	0.52
lose-change	A	0.39	0.38
lose-change	B	0.47	0.38

Table 5: rates of accordance to win-stay lose-change strategy (averages over all subjects in non-coordinating pairs)

In fact, rather the opposite can be found. All categories show a downward trend from the first half to the second half of the experiment, though some trends are only small in size. We may, hence, conclude that gradual adjustment to the win-stay lose-change strategy can neither be observed. However, wehave no account for the general trend to move away from both the win-stay part and the lose-change part of the strategy.

Using maximum-likelihood estimation we calculated the best fitting parameters for both learning schemes. We did separate analyses for the full set of pairs and for the set of pairs that do not converge because *ad hoc* there is no reason why the learning schemes should only be applied to predict non-converging pairs. We estimated the learning schemes over all 100 rounds.

	All pairs			non-converging pairs		
restrictions	$p_A(1) = p_B(1) = 1$	$p_A(1) = p_B(1)$	None	$p_A(1) = p_B(1) = 1$	$p_A(1) = p_B(1)$	None
$p_A(1)$			1.53			8.32
$p_B(1)$		1.97	2.43		6.92	7.50
j	0.984	0.976	0.977	1.0025	0.984	0.983
-2 LL	3361.94	3345.85	3328.49	2688.48	2626.18	2625.04

Table 6: maximum-likelihood estimation for reinforcement learning

[18] A detailed table on an individual level is given in Appendix 7.3.

[19] In actual fact, the evidence of Rabinowitz et al.'s table 2 in favour of the win-stay lose-change strategy is rather weak and mainly concerns the win-stay part.

	all pairs				non-converging pairs			
	m = 1	*m = 2*	*m = 3*	*m = 4*	*m = 1*	*m = 2*	*m = 3*	*m = 4*
X	0.83	0.84	0.90	0.92	0.64	0.60	0.60	0.57
-2LL	3732.34	3798.85	3697.68	3657.80	2707.34	2746.49	2759.58	2768.66

Table 7: maximum likelihood estimation for experimentation learning

The analysis shows for the reinforcement learning scheme that taking all pairs into account we get results very close to our initial assumptions, i.e. $p_A(1)$ and $p_B(1)$ at approximately 1 and j very near 1, allowing for only 2 to 2.5% forgetting. Exclusion of the converging pairs yields similar values for j but much higher values for the initialisation parameters. This is plausible since higher initial propensities mean more variation of play. Consequently, fixing the initial parameter at 1 yields an implausible value for j[20]. Not surprisingly, if considering all pairs fixing the initial propensities to be equal yields a significant loss of explanatory power. The propensity of B naturally is much larger than the one of A. Interestingly, when excluding the converging pairs this restriction does not have a significant impact. This is one more piece of evidence in favour of the hypothesis that those subjects who do not coordinate with their partners have no systematic tendency towards either of the actions.

For the experimentation learning we first find that when excluding the converging pairs m = 1 again fits the data best. The small values for x indicate that the win-stay part of the strategy is not very pronounced. Considering all pairs, values for x are naturally much higher since at the point of convergence the win-stay part is fully met. As for the performance, it might seem surprising that m = 4 is best and most probably higher values for m would yield even better results. The reason is simply that because of the value for x being predicted much higher those strategies perform better which allow for more random play in pairs that do not converge.

A comparison between the two learning schemes shows that we can add our study to the list of those that favour reinforcement learning. In order to put the comparison on equal footing we compare the best likelihood value from the experimentation learning scheme with the likelihood value for the one-parameter reinforcement model. This shows that for both categories the reinforcement model performs much better than the experimentation learning scheme. This is so, even though the experimentation learning scheme has the advantage of one more discrete variable (namely m). However, we have to point out that qualitative features force us to say that none of the algorithms does a particularly good job.

[20] $j > 1$ means that experience further in the past have a stronger impact on current action than recent experiences.

4.2 Game space

One may argue that, since people do not know which game they are actually going to play, they will simply imagine one and start playing according to that game. However, this does not cause trouble to our interpretations of subjects' play, because, given subjects' prior information, there are only four games that can possibly be played. In particular, if one writes down all 256 possible 2x2 games resulting from payoffs that can take either value 0 or 1 and subsequently eliminates all games that are either trivial or involve a dominant strategy for at least one player, then you are left with one of the following games: (i) the matching pennies game, (ii) the symmetric coordination game, (iii) our mutual fate control game, or (iv) fate control – behaviour control[21], in which, for one player, giving is identified with one strategy and the other player's giving strategy depends on the strategy chosen by the first player. For all cases experimentation learning eventually leads either to the unique mixed-strategy Nash-equilibrium or to an outcome that maximizes both, individual as well as joint payoffs.

Furthermore, the only game not containing the prospect of eventual coordination is the game of matching pennies[22]. Mookherjee and Sopher (1994) already conducted this game under little information and found little evidence for deliberate mixing behaviour. We investigated whether subjects in our experiment exhibited similar patterns of behaviour.

We first checked for the consistency of mixing between the two strategies. Assuming that, in the first periods, people learn about their environment and, after a sufficient amount of time, they converge to a more or less stable mixed strategy, we expect to observe a mix of strategies that conforms to an i.i.d. play. Taking the observed rate of occurrence of strategy B to be the approximation to the probability of playing B, one can conduct a Wald-Wolfowitz One-Sample Runs test to check whether the number of runs (i.e. the number of sequences in which the same strategy has been played) is similar to the expected number of runs under the i.i.d. hypothesis. We perform this test for the last 40 rounds of each player and exclude players of pairs that do coordinate. The results are shown in Table 8.

[21] For a more detailed description of fate control – behaviour control, see e.g. Arickx and van Avermaet (1981).
[22] Note that the crucial difference in dynamics between mutual fate control and matching pennies lies in the impact of opponent's play. In matching pennies the opponent's action determines *which of my own strategies* is vindicated and which is refuted, while in mutual fate control the opponent's action determines *whether my currently active strategy* is vindicated or refuted.

player	Freq B	No. of runs	exp. no. of runs	p-value	
1	22	27	20,80	0,05	*
2	23	19	20,55	0,63	
7	33	15	12,55	0,30	
8	20	40	21,00	0,00	**
9	21	19	20,95	0,63	
10	21	20	20,95	0,88	
11	18	11	20,80	0,00	**
12	19	19	20,95	0,63	
13	3	3	6,55	0,00	**
14	19	7	20,95	0,00	**
15	13	19	18,55	1,00	
16	11	14	16,95	0,30	
21	18	17	20,80	0,26	
22	21	26	20,95	0,11	
23	2	3	4,80	0,05	*
24	17	11	20,55	0,00	**
27	21	27	20,95	0,04	**
28	16	8	20,20	1,00	
29	22	13	20,80	1,00	
30	32	9	13,80	0,99	

Table 8: Wald-Wolfowotz One-Sample Runs test for last 40 periods
(* 10% significance, ** 5% significance)

As can be seen, for 8 out of 20 subjects the i.i.d. hypothesis can be rejected to the 10%-level. For the 5%-level the number of rejections is still 6 out of 20. This is in line with what Mookherjee and Sopher (1994) have found for their low-information treatment. In our study 5 out of 8 rejections are based on too few runs, which is not an immediate indicator for some sort of inertia in subject's play. But, two of the rejections that are caused by too many runs (players 8 and 27) are due to the players obviously alternating between the two strategies. For further analysis of inertia in subject's play we have to look at the individual dynamics.

For this purpose we follow Mookherjee and Sopher (1994) and adopt the concept of vindication of strategies, which has first been used for the definition of basic routine learning schemes by Bush and Mosteller (1955). According to them, a strategy is vindicated whenever either, it has been chosen and yielded a success , or, the alternative strategy has been chosen and yielded a failure. For the following, we denote the strategy chosen by player i at time t again by $a^i(t)$ and the resulting payoff by $p^i(t)$. The vindication of player i's strategy 1 will be denoted by $v^i(t)$, thus,

$$v^i(t) = \begin{cases} 1 & \text{if } \left[\left(a^i(t)=1\right) \wedge \left(p^i(t)=1\right)\right] \vee \left[\left(a^i(t)=0\right) \wedge \left(p^i(t)=0\right)\right] \\ 0 & \text{otherwise} \end{cases}.$$

We use a logistic regression to analyse the impact of own lagged actions and lagged vindication of strategy B (up to four periods before the current period) on the probability to

choose strategy B, the rewarding strategy. Table 9 shows the coefficients of the respective regressands and the corresponding p-values. We excluded groups that eventually co-ordinate on the efficient outcome.

	coefficient	p-value
constant	-1.26	0.00
a(t-1)	0.80	0.00
a(t-2)	0.52	0.00
a(t-3)	-0.27	0.01
a(t-4)	0.69	0.00
v(t-1)	0.04	0.73
v(t-2)	0.31	0.00
v(t-3)	-0.37	0.00
v(t-4)	0.57	0.00

Table 9: logistic regression on probability of choosing strategy B for last 40 rounds (n = 1920, $\chi^2(8) = 198.74$, p = 0.00)

Large positive coefficients and strong significance for the one- and two-period lagged own choices signify a strong tendency towards inertia. However, three-lagged own choices are significantly negative. Furthermore, vindication is not immediately responded to as can be seen from the insignificant coefficient for the one-period lagged vindication. Former vindication, however, does have a significant impact, though with alternating sign. We conclude that subjects exhibit a significant amount of inertia and respond volatile to the vindication of the rewarding strategy. We find positive auto-correlation of subject's choices. In this respect behaviour differs from mainly negatively auto-correlated choices in the matching-pennies game. However, the volatile response to vindication is a similarity found in both games under low information conditions. Neither reinforcement learning nor experimentation learning can account for this latter observation.

4.3 Individual patterns

Under little information it is likely that subjects exhibit varying strategies in order to explore their environment. Furthermore, the task looks similar to playing a two-armed bandit, so it is likely that subjects use individual exploration strategies before playing a certain action. As noted earlier, some of the subjects deliberately chose to play certain patterns. In the following table we give an overview of broad characteristics of the subjects.

player	**Characteristic**
1	alternating play in rounds 11 to 48 and 50 to 73
2	some long sequences of the same strategy until round 45, no pattern thereafter
3	co-ordination starting in round 2
4	co-ordination starting in round 2
5	alternating play in rounds 32 to 45, co-ordination starting in round 65

6	co-ordination starting in roun d 65
7	many long sequences of the same strategy, in later rounds more sequences of Bs
8	sequence of As in rounds 6 to 30, alternating play from round 46 onwards
9	no specific pattern
10	a few short sequences of small patterns like (AABBAA) or (ABA)
11	some long sequences of the same strategy at the beginning and towards the end, alternating play in rounds 46 to 60
12	no specific pattern
13	very long sequences of the same strategy at the beginning and during the whole second half, alternating play in rounds 18 to 40
14	many long sequences of the same strategy throughout
15	less changes than predicted by independent draws, though no specific pattern
16	one larger sequence of As towards the end, otherwise no specific pattern
17	co-ordination starting at round 4
18	co-ordination starting at round 4
19	co-ordination starting at round 2
20	co-ordination starting at round 2, deviation in the last 2 rounds
21	less changes than predicted by independent draws, though no specific pattern, more changes towards the end
22	no specific pattern
23	almost always As
24	many long sequences of the same strategy
25	coordination starting in round 2
26	coordination starting in round 2
27	less changes than predicted by independent draws, though no specific patter n, more changes towards the end
28	many long sequences of the same strategy
29	no specific pattern
30	some alternation patterns in the intermediate term, rather long sequences towards the end
31	no pattern at the beginning, almost always Bs starting in round 53
32	no pattern at the beginning, almost always Bs starting in round 53
33	coordination starting in round 9
34	coordination starting in round 9

Table 10: Characteristics of subjects' play

As can be seen from Table 10, only 4 out of the 20 subjects that do not succeed in co-ordinating with their partner do not show any sign of inertia or a deliberate application of a dynamic pattern. By the use of these patterns, subjects obviously try to check out their environment. The most popular patterns are long sequences of the same strategy (10 subjects) and the alternation between the two strategies (6 subjects). There is very little evidence in

favour of gradual adjustment towards strategy B. Only one subject (player 7) can be categorised into this class. Players 31 and 32 (constituting the converging pair 16) rather suddenly switch to almost always play B.

Most striking is the persistent play of strategy A by player 23. This subject obviously was not willing to deviate from his favourite action even after long sequences of non-rewarding feedback. Only twice he felt driven to play action B for two times before falling back to play A. In the questionnaires, that were handed out after the completion of the 100 rounds and before subjects were paid off, this subject wrote: "Although I always tried to choose the same strategy I did not reach to always get a payoff of 1; possibly because the other [player] misinterpreted the situation and chose this and that." The remark clearly indicates that this person had a pure coordination game in mind and did not think of any other possible payoff scheme. In fact, 5 out of the 20 subjects that did not coordinate with their partner stated that they had the impression they were playing a coordination game. Most of the remaining 15 players stated that they did not have any clue as to what the payoff matrix must have looked like or indicated a repeating multi-period scheme that should have given them the desired payoff of 1, despite our care to indicate in the instruction sheet that payoffs only depended on the decisions of the stage game. Several subjects also indicated that by using certain patterns they tried to explore the payoff environment and the way the opponent was playing.

It seems that within an environment with little information subjects are tempted to believe they are involved in a single-person experiment or to focus on a coordination task. However, none of these believes justifies the renunciation of the experimentation learning scheme.

5. Conclusion

The present experiment was designed to explore learning behaviour under little information. Subjects were not given the full payoff matrix; instead they were given the amount of information that is just sufficient to perform simple adaptive learning. The game played was mutual fate control where each of the players decides upon giving reward to the opponent or not. Two intuitively plausible learning schemes are presented which make distinct qualitative predictions. The reinforcement model predicts that people slowly adjust propensities, and, after a large number of periods, profiles converge to any of a diffuse set of possible profiles. To the contrary, the experimentation learning makes the strong prediction that people will exhibit some random behaviour at the beginning of the game and, with high probability, pairs will converge to the efficient outcome within a limited amount of time.

The observations stand in sharp contrast to results from previous studies of this game that have mainly been run by psychologists. Little coordination can be observed, and convergence

appears to be triggered by extreme chance events. Moreover, compliance to simple dynamics, such as the win-stay lose-change scheme, cannot be found. Our impression is that even the simultaneous-move environment that Kelley et al. (1962) consider as particularly favourable for coordination is still not sufficient to really enforce this learning dynamic. There are two events that have to coincide in order to trigger coordination. First, both players have to choose B simultaneously, and second, none of the players deviates in the next interaction. If one player deviates, an immediate vindication of A follows. Without a device that coordinates timing this task seems to be difficult to accomplish. Furthermore, the more flexible experimentation learning scheme leads to the desired result only if one may expect the partner to accord to this dynamic, too. As we have seen from the data, there is good reason for subjects to doubt this.

Generally, the results are unfavourable for all learning schemes presented so far. Convergence to the efficient outcome is much slower than can reasonably be traced by the experimentation learning scheme. The basic win-stay lose-change scheme is not accorded to, either. Furthermore, subjects do not show any sign of slow adjustment behaviour. Rather, subjects seem to follow an exploratory scheme, much in the way Kalai and Lehrer (1995) describe it by using the concept of the environment response function. Subjects use different multi-period patterns in order to explore the response of their individual environment. However, Table 10 clearly shows that they obviously do not conform to any simple dynamics and stick to complex exploration schemes even if they do not lead to a significant improvement in payoff. Hence, mutual fate control provides an environment, in which subjects' behaviour is more complex than it should be. Such behaviour results in avoidable inefficiencies.

The findings of this study pose new questions that should be addressed in future research. First, we have seen that to withhold any information on the payoff scheme generally deprives subjects of the ability to systematically coordinate on the efficient cell, even though, theoretically, there exists a simple learning scheme that does the job. Had subjects known the payoff matrix, without doubt they would have coordinated instantly. Hence, there must be a certain amount of information in-between that just suffices for people to coordinate. Which one, is still an open question. Second, more research is needed to successfully describe how individuals explore their environment. This experiment supports the view that individuals do not simply update propensities of choices, they rather combine experimentation with some heuristics in order to explore the environment. Only recently, there have some attempts been made to capture exploration behaviour. Most notably is the concept of the environment

response function by Kalai and Lehrer (1995). They argue that in low -information settings people tend to consider their feedback as a composite of the whole environment they are acting in. Unfortunately, the authors do not specify the way subjects may perform their updating of the response function. It is a future task to use this theory to broaden the set of empirically supported learning rules.

6. References

Arickx, M. and Van Avermaet, E. (1981) "Interdependent Learning in a Minimal Social Situation" *Behavioral Science* 26, 229 – 242.

Börgers, Tilman and Sarin, Rajiv (1997) "Learning Through Reinforcement and Replicator Dynamics" *Journal of Economic Theory* 77, 1 – 14.

Bosch-Domènech, Antoni and Vriend, Nicolaas (1999) "Imitation of Successful Behavior in Cournot Markets" working paper, Queen Mary and Westfield College, University of London, UK

Brown, George W. (1951) "Iterative Solutions of Games by Fictitious Play" in "Activity Analysis of Production and Allocation", ed. T. C. Koopmans, New York, Wiley

Bush, Robert R. and Mosteller, Frederick (1955) "Stochastic Models for Learning" New York, Wiley.

Camerer, Colin and Ho, Teck-Hua (1999) "Experienced-Weighted Attraction Learning in Normal Form Games" *Econometrica* 67, 827 – 874.

Cheung, Yin-Wong and Friedman, Daniel (1997) "Individual Learning in Normal Form Games: Some Laboratory Results" *Games and Economic Behavior* 19, 46 – 76.

Conlisk, John (1993) "Adaptive tactics in games – Further solutions to the Crawford puzzle" *Journal of Economic Behavior and Organization* 22, 51 – 68.

Crawford, Vincent (1995) "Adaptive Dynamics in Coordination Games" *Econometrica* 63, 103 – 144.

Cross, John G. (1973) "A Stochastic Learning Model of Economic Behavior" *Quarterly Journal of Economics* 87, 239 – 266.

Duffy, John and Feltovich, Nick (1997) "Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study" working paper, Department of Economics, University of Pittsburgh.

Erev, Ido and Rapoport, Amnon (1998) "Coordination, "Magic," and Reinforcement Learning in a Market Entry Game" *Games and Economic Behavior* 23, 146 – 175.

Erev and Roth (1998) "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria" *American Economic Review* 88 (4), 848 – 881.

Fudenberg, Drew and Levine, David (1998) "The Theory of Learning in Games" MIT Press.

Hon-Snir, Shlomit, Monderer, Dov, and Sela, Aner (1998) "A Learning Approach to Auctions" *Journal of Economic Theory* 82, 65 – 88.

Huck, Steffen, Normann, Hans-Theo, and Oechssler, Jörg (1998) "Does Information About Competitors' Actions Increase or Decrease Competition in Experimental Markets?" working paper, Humboldt-University of Berlin.

Huck, Steffen, Normann, Hans-Theo, and Oechssler, Jörg (1999) "Learnin in Cournot Oligopoly – An Experiment" *Economic Journal* 109, C80 – C95.

Kalai, Ehud and Lehrer, Ehud (1995) "Subjective Games and Equilibria" *Games and Economic Behavior* 8, 123 – 163.

Kelley, H.H., Thibaut, J.W., Radloff, R., Mundy, D. (1962) "The Development of Cooperation in the 'Minimal Social Situation'" *Psychological Monographs* 76, whole no. 19.

Milgrom, Paul and Roberts, John (1990) "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities" *Econometrica* 58 (6), 1255 – 1277.

Milgrom, Paul and Roberts, John (1991) "Adaptive and Sophisticated Learning in Normal Form Games" *Games and Economic Behavior* 3, 82 – 100.

Mookherjee, Dilip and Sopher, Barry (1994) "Learning Behavior in an Experimental Matching Pennies Game" *Games and Economic Behavior* 7, 62 – 91.

Mookherjee, Dilip and Sopher, Barry (1997) "Learning and Decision Costs in Experimental Constant Sum Games" *Games and Economic Behavior* 19, 97 – 132.

Nagel, Rosemarie and Vriend, Nicolaas J. (1997) "An Experimental Study of Adaptive Behavior in an Oligopolistic Market Game" working paper, Universtitat Pompeu Fabra, Barcelona, Spain.

Rabinowitz, L., Kelley, H.H., and Rosenblatt, R.M. (1966) "Effects of Different Types of Interdependence and Response Conditions in the Minimal Social Si tuation" *Journal of Experimental Social Psychology* 2, 169 – 197.

Robinson, Julia (1951) "An Iterative Method of Solving a Game" *Annals of Mathematics* 54 (2), 296 – 301.

Roth, Alvin E. and Erev, Ido (1995) "Learning in Extensive -Form Games: Experimental D ata and Simple Dynamic Models in the Intermediate Term" *Games and Economic Behavior* 8, 164 – 212.

Sidowski, J.B. (1957) "Reward and Punishment in the Minimal Social Situation" *Journal of Experimental Psychology* 54, 318 – 326.

Sidowski, J.B., Wyckoff, L.B., and Tabory, L. (1956) "The Influence of Reinforcement and Punishment in a Minimal Social Situation" *Journal of Abnormal and Social Psychology* 52, 115 – 119.

Smith, Vernon L. (1990) "Experimental Economics: Behavioral Lessons for Microeconomic Theory and Policy" Nancy L. Schwartz Memorial Lecture, Northwestern University.

Tang, Fang-Fang (1998) "A Comparative Study on Learning and Stability in Normal Form Games: Some Experimental and Simulational Results" working paper, Department of Economics, National University of Singapore.

Thibaut, John W. and Kelley, Harold H. (1959) "The Social Psychology of Groups" New York, Wiley.

Van Huyck, John B., Battalio, Raymond C., and Rankin, Frederick W. (1996) "Selection Dynamics and Adaptive Behavior Without Much In foration" working paper, Texas A&M University.

7. Appendices

7.1 Instructions

Preliminaries:

You are participating in a study on decision making within the framework of experimental economics. After you have read these instructions we will come to you in ord er to clarify open questions. If you encounter any further questions during play please raise your hand, so a member of the staff can come to you.

Please do not touch the computer until you are asked to start the game. Tips on the usage of the computer are given on the back page.

During the session you will face a sequence of decision tasks. At each decision you may earn money. How much money you will get depends on your decisions. The aggregated payoff will be paid to you in cash at the end of the experi ment. Your decisions as well as your payoff

is known only to you, that is we take care that no other participant will get any information on your decisions or payoffs.

Decision task:

You are one of 8 (10) persons. All persons are faced with the same task, that is, all persons have got identical instruction sheets and sit, separated from each other, in front of a computer terminal.

Before the start of the first round, all 8 (10) persons are matched randomly into four (five) pairs. The pairs remain fixed throughout the whole session, that is, the person you are playing with will be the same throughout the whole session. You will never be told whom you are paired with.

Soon, you will be required to make the following identical decision in 100 successive rounds:

In each round you have got the choice between two alternatives: action A and action B. After you and the person you are paired with have entered the decisions upon your choice the computer will calculate your payoff according to a predetermined scheme and subsequently informs you about your payoff.

Payoff scheme:

The payoff scheme will not be made public. What is known is that the payoffs within your pair solely depend on your decision and the decision of your opponent of that round. Particularly, no random process will be used to calculate payoffs. Furthermore, it is known that payoffs can take either of the values, 0 or 1, where 0 is equivalent to 0 DM and 1 is equivalent to 0.30 DM.

Your MAXLAB team

Tips for the usage of the computer:

1. Please do not touch the computer until you are asked to start the play.
2. Please do not start any other programs during the experimental session.
3. You can enter your decision by clicking with your mouse on one of the two buttons, A or B. After that, a window appears that asks you to confirm your choice.
4. After confirming your choice, please wait until the information on the last round appears before you go on with the next round.
5. In case there are any technical problems with the usage of the program, please immediately inform a member of the staff by raising your hand.

7.2 Data

round	pair 1		pair 2		pair 3		pair 4		pair 5		pair 6		pair 7		pair 8		pair 9		pair 10		pair 11		pair 12		pair 13		pair 14		pair 15		pair 16		pair 17	
player	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	A	A	A	A	A	B	B	A	B	A	B	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A	B	A	B	A	A	B	A	B
2	A	B	B	B	A	A	A	A	A	B	B	B	B	A	B	A	A	A	B	B	B	A	A	B	B	A	B	B	A	A	B	B	A	B
3	A	B	B	B	A	A	A	A	A	A	B	A	B	A	B	A	A	A	B	B	A	B	A	B	B	B	B	B	A	A	B	B	A	B
4	A	A	B	B	B	B	A	B	B	A	B	B	B	A	A	A	B	B	B	B	A	B	A	A	B	B	B	A	A	B	B	B	A	B
5	A	A	B	B	A	B	A	B	B	A	B	B	A	A	B	B	B	B	B	B	B	A	A	B	B	B	B	A	B	B	A	B	A	B
6	A	A	B	B	B	B	A	A	A	A	B	B	A	A	B	B	B	B	B	B	A	A	A	B	B	B	B	B	A	B	B	A	A	B
7	B	B	B	B	A	B	B	A	B	B	B	A	A	A	B	A	B	B	B	B	A	B	A	B	B	B	B	B	B	A	B	A	A	A
8	A	B	B	B	B	A	B	A	A	A	B	A	A	B	A	A	B	B	B	B	A	B	A	B	B	B	B	A	B	B	A	A	A	A
9	A	A	B	B	B	B	A	A	B	A	A	A	A	B	A	B	B	B	B	B	A	B	A	A	B	B	B	A	A	A	B	B	B	B
10	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	A	A	B	B	B	B	A	A	A	B	B	B
11	B	B	B	B	A	B	A	A	A	B	B	A	A	A	A	A	B	B	B	B	B	A	A	B	B	B	B	B	A	B	A	A	B	B
12	A	B	B	B	A	B	B	A	B	B	B	A	A	A	B	A	B	B	B	B	B	B	A	B	B	B	A	A	B	A	B	B	B	B
13	B	B	B	B	A	A	B	A	A	A	B	A	A	A	B	B	B	B	B	B	A	A	A	B	B	B	A	A	B	B	A	B	B	B
14	A	B	B	B	B	A	B	A	A	A	A	B	B	A	B	A	B	B	B	B	A	B	A	A	B	B	A	A	B	A	B	A	B	B
15	B	B	B	B	B	A	A	A	A	A	A	B	B	A	A	B	B	B	B	B	B	A	A	A	B	B	A	B	A	B	A	A	B	B
16	A	B	B	B	A	A	A	A	B	B	A	B	B	A	A	A	B	B	B	B	A	B	A	B	B	B	A	A	A	A	B	B	B	B
17	B	B	B	B	B	B	A	A	A	B	A	A	B	A	A	A	B	B	B	B	A	B	A	A	B	B	A	A	B	A	A	A	B	B
18	A	B	B	B	A	A	A	A	B	A	A	B	B	A	A	B	B	B	B	B	A	A	A	B	B	B	A	A	B	B	B	A	B	B
19	B	B	B	B	B	A	A	A	B	A	A	B	A	A	A	B	B	B	B	B	B	A	A	B	B	B	A	A	A	B	B	A	B	B
20	A	A	B	B	B	A	B	A	B	B	A	A	B	B	A	B	B	B	B	B	A	A	A	B	B	B	A	B	B	B	B	A	B	B
21	B	B	B	B	B	B	B	A	A	B	A	A	A	B	B	A	B	B	B	B	A	A	A	B	B	B	A	A	A	A	A	A	B	B
22	A	B	B	B	B	B	B	A	A	A	B	A	B	A	A	B	B	B	B	B	A	B	A	B	B	B	A	B	B	B	B	B	B	B
23	B	B	B	B	B	A	B	A	A	A	B	A	A	A	A	A	B	B	B	B	B	A	A	A	B	B	B	A	B	B	A	B	B	B
24	A	B	B	B	A	A	B	A	B	B	B	B	B	B	A	B	B	B	B	B	B	A	A	B	B	B	B	A	B	A	B	A	B	B
25	B	B	B	B	A	A	B	A	A	A	B	B	A	A	A	A	B	B	B	B	A	A	A	B	B	B	B	A	A	B	B	A	B	B
26	A	A	B	B	A	B	B	A	B	B	B	B	B	A	A	A	B	B	B	B	A	A	A	A	B	B	B	A	A	B	A	A	B	B
27	B	B	B	B	A	A	A	A	B	A	B	A	A	A	B	B	B	B	B	B	A	A	A	A	B	B	B	A	A	A	A	B	B	B
28	A	B	B	B	B	A	A	A	A	B	B	A	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	A	A	A	A	B	B	B
29	B	A	B	B	B	A	A	A	A	A	A	A	A	B	B	A	B	B	B	B	A	B	A	B	B	B	A	A	A	B	B	A	B	B
30	A	B	B	B	A	A	A	A	B	A	A	B	B	A	A	A	B	B	B	B	B	A	A	B	B	B	A	B	A	A	A	A	B	B
31	B	A	B	B	A	B	A	B	B	B	A	A	A	B	B	B	B	B	B	B	B	A	A	B	B	B	B	A	A	B	B	B	B	B
32	A	A	B	B	B	B	A	B	B	B	A	A	B	B	B	B	B	B	B	B	B	A	A	A	B	B	B	A	A	B	B	B	B	B
33	B	B	B	B	A	B	A	A	B	A	B	B	A	A	B	B	B	B	B	B	A	B	A	A	B	B	A	A	A	A	A	B	B	B
34	A	B	B	B	B	A	A	B	A	B	B	A	B	A	A	B	B	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B	B
35	B	B	B	B	A	B	A	B	B	B	B	A	A	A	A	B	B	B	B	B	B	A	B	A	B	B	B	B	B	B	B	A	B	B
36	A	B	B	B	B	A	A	A	A	B	A	B	B	A	B	A	B	B	B	B	B	A	B	A	B	B	A	B	B	A	A	A	B	B
37	B	B	B	B	A	A	A	A	A	A	B	B	A	A	B	A	B	B	B	B	B	A	A	B	B	B	A	B	B	B	A	B	B	B
38	A	B	B	B	B	B	A	A	A	B	A	B	B	A	B	A	B	B	B	B	B	B	A	B	B	B	B	A	B	B	A	B	B	B
39	B	B	B	B	A	B	B	A	B	A	B	A	A	A	B	B	B	B	B	B	A	A	A	A	B	B	B	A	A	B	A	A	B	B
40	A	A	B	B	B	A	A	A	B	A	B	A	B	A	A	A	B	B	B	B	A	A	A	A	B	B	A	A	B	A	B	A	B	B
41	B	A	B	B	A	B	A	B	B	B	A	A	B	A	B	B	B	B	B	B	A	B	A	B	B	B	A	B	B	B	B	A	B	B
42	A	A	B	B	B	B	A	B	B	A	A	A	A	A	A	B	B	B	B	B	A	B	A	B	B	B	A	B	B	B	B	A	B	B
43	B	A	B	B	A	B	A	A	A	B	A	B	A	A	A	A	B	B	B	B	A	B	A	B	B	B	A	A	A	A	A	A	B	B
44	A	A	B	B	B	A	A	A	A	B	B	B	B	A	B	A	B	B	B	B	A	B	A	B	B	B	A	B	A	B	A	B	B	B
45	B	A	B	B	A	A	A	A	B	A	B	B	B	A	B	A	B	B	B	B	B	A	A	A	B	B	A	A	A	B	A	B	B	B
46	B	B	B	B	A	B	A	A	A	A	B	A	A	A	A	A	B	B	B	B	B	A	A	A	B	B	A	A	A	A	A	A	B	B
47	A	B	B	B	B	A	B	B	B	B	A	B	A	A	B	B	B	B	B	B	B	A	A	A	B	B	B	A	B	A	A	A	B	B
48	B	A	B	B	B	A	A	A	A	A	B	B	B	A	B	A	B	B	B	B	B	A	A	A	B	B	B	A	B	A	A	B	B	B
49	B	B	B	B	A	B	B	B	A	B	A	B	A	A	A	B	B	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B	B
50	B	A	B	B	B	A	B	A	B	A	B	B	B	A	A	B	B	B	B	B	A	B	A	A	B	B	A	B	B	B	B	A	B	B
51	A	A	B	B	B	A	B	B	B	A	A	B	A	A	A	B	B	B	B	B	A	B	A	A	B	B	B	A	B	A	A	A	B	B
52	B	B	B	B	A	A	A	A	B	A	B	A	A	A	B	A	B	B	B	B	A	A	A	B	B	B	A	A	A	A	A	A	B	B
53	A	A	B	B	B	B	B	B	B	A	A	B	A	A	B	A	B	B	B	B	B	B	A	B	B	B	A	B	A	B	B	B	B	B
54	B	A	B	B	A	B	B	A	B	B	B	B	A	A	A	B	B	B	B	B	B	A	A	B	B	B	A	A	A	A	B	B	B	B
55	A	B	B	B	B	A	B	B	B	B	A	B	A	A	A	A	B	B	B	B	B	B	A	B	B	B	A	A	B	B	A	A	B	B
56	B	A	B	B	B	A	B	A	A	B	B	B	A	A	B	B	B	B	B	B	B	A	A	B	B	B	B	A	B	B	B	B	B	B
57	A	B	B	B	B	B	B	A	A	A	A	A	A	B	B	B	B	B	B	B	B	A	A	B	B	B	B	A	B	B	B	B	B	B
58	B	B	B	B	B	A	B	A	A	B	B	A	B	B	B	B	B	B	B	B	B	B	A	A	B	B	B	A	A	B	B	B	B	B
59	A	A	B	B	A	A	B	B	B	A	A	A	A	B	A	B	B	B	B	B	A	A	A	B	B	B	B	A	A	B	B	B	B	B
60	B	B	B	B	A	A	A	A	A	A	B	A	A	B	A	B	B	B	B	B	A	A	A	B	B	B	B	A	B	B	A	B	B	B
61	A	B	B	B	B	B	B	B	B	B	B	A	A	B	A	A	B	B	B	B	A	B	A	A	B	B	B	A	B	B	A	A	B	B
62	B	A	B	B	A	A	A	A	B	B	B	A	A	B	B	A	B	B	B	B	A	A	A	B	B	B	A	A	A	B	B	B	B	B
63	A	B	B	B	B	B	B	B	A	B	A	A	A	A	A	B	B	B	B	B	B	B	A	B	B	B	A	B	A	B	B	B	B	B
64	B	B	B	B	A	A	B	A	A	A	A	B	A	A	A	A	B	B	B	B	B	B	A	B	B	B	B	A	B	B	B	B	B	B
65	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	B	B	B	A	B	B	B	B	B	B	B
66	B	A	B	B	B	B	B	A	A	B	A	A	A	A	A	A	B	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B	B
67	A	B	B	B	B	B	B	B	B	A	B	A	A	A	A	B	B	B	B	B	A	B	A	A	B	B	A	B	B	B	B	B	B	B
68	B	B	B	B	B	B	A	A	B	A	A	A	A	A	A	B	B	B	B	B	A	B	A	A	B	B	B	A	A	B	B	B	B	B
69	A	B	B	B	B	B	B	B	A	A	A	B	A	B	A	A	B	B	B	B	B	A	A	B	B	B	B	A	A	B	B	B	B	B
70	B	B	B	B	B	B	B	A	B	A	A	B	A	B	B	A	B	B	B	B	B	B	A	B	B	B	A	A	B	B	A	B	B	B
71	A	A	B	B	B	B	B	B	B	A	A	B	A	B	B	B	B	B	B	B	B	A	A	B	B	B	A	B	B	B	B	A	B	B
72	B	A	B	B	B	B	A	A	B	A	A	B	A	B	B	A	B	B	B	B	B	B	A	B	B	B	B	B	A	A	B	B	B	B
73	A	A	B	B	B	B	B	B	A	A	A	B	A	B	B	A	B	B	B	B	A	A	A	B	B	B	A	B	A	A	B	B	B	B
74	A	B	B	B	B	B	B	A	B	B	A	A	A	B	A	A	B	B	B	B	A	A	A	B	B	B	B	B	A	B	B	B	B	B
75	A	B	B	B	B	B	B	B	B	B	A	B	A	B	A	B	B	B	B	B	A	A	A	B	B	B	A	B	B	B	B	B	B	B
76	B	A	B	B	B	B	B	A	A	A	A	B	B	A	B	A	B	B	B	B	A	B	A	B	B	B	B	B	B	B	B	B	B	B
77	B	A	B	B	B	B	B	B	B	A	A	A	B	A	A	B	B	B	B	B	A	A	A	A	B	B	A	B	B	B	B	B	B	B
78	B	A	B	B	B	B	A	A	A	B	A	A	A	B	B	B	B	B	B	B	A	B	A	A	B	B	B	B	A	A	B	B	B	B
79	A	A	B	B	B	B	B	B	A	B	B	B	A	A	A	B	B	B	B	B	B	B	A	A	B	B	A	A	A	B	B	B	B	B
80	B	B	B	B	B	B	B	A	B	A	B	B	A	A	B	A	B	B	B	B	A	A	A	A	B	B	B	A	A	B	A	B	B	B
81	A	B	B	B	B	B	B	B	B	A	B	B	A	A	A	A	B	B	B	B	A	A	A	A	B	B	B	A	B	B	A	A	B	B
82	B	B	B	B	B	B	B	A	A	B	B	A	A	A	A	A	B	B	B	B	A	B	A	A	B	B	B	A	B	B	B	B	B	B
83	A	A	B	B	B	B	B	B	B	B	A	B	A	A	A	A	B	B	B	B	B	B	B	A	B	B	B	A	B	B	B	B	B	B
84	B	A	B	B	B	B	B	A	B	B	A	B	A	B	B	A	B	B	B	B	B	B	B	A	B	B	B	A	B	A	B	A	B	B
85	B	B	B	B	B	B	B	B	B	A	A	A	A	B	A	B	B	B	B	B	A	A	A	A	B	B	A	A	A	A	B	B	B	B
86	B	B	B	B	B	B	A	A	B	B	B	A	B	A	A	A	B	B	B	B	A	B	A	A	B	B	A	B	B	A	B	B	B	B
87	B	A	B	B	B	B	B	B	A	A	B	B	A	B	A	A	B	B	B	B	A	A	A	A	B	B	A	B	B	B	B	B	B	B
88	A	A	B	B	B	B	B	A	A	B	B	A	A	B	A	A	B	B	B	B	B	A	A	B	B	B	A	A	B	B	B	B	B	B
89	B	B	B	B	B	B	B	B	A	A	A	A	A	B	A	A	B	B	B	B	B	B	A	B	B	B	B	A	B	B	B	B	B	B
90	A	A	B	B	B	B	A	A	A	B	A	B	A	A	A	A	B	B	B	B	B	A	A	A	B	B	A	A	B	A	A	B	B	B
91	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	B	B	A	B	A	A	B	B	B	A	A	A	A	B	B	B
92	B	B	B	B	B	B	B	A	A	B	B	A	A	A	B	A	B	B	B	B	B	A	A	B	B	B	B	A	A	A	B	A	B	B
93	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	B	B	B	B	B	A	A	A	B	B	A	A	A	B	B	A	B	B
94	A	A	B	B	B	B	B	A	A	A	A	A	A	A	B	A	B	B	B	B	A	B	A	A	B	B	B	A	A	B	A	A	B	B
95	B	A	B	B	B	B	B	B	A	B	B	A	A	A	B	A	B	B	B	B	B	B	A	A	B	B	B	A	A	B	B	B	B	B
96	B	B	B	B	B	B	A	A	B	B	B	A	A	B	A	A	B	B	B	B	B	A	A	A	B	B	A	A	A	B	B	B	B	B
97	B	B	B	B	B	B	B	B	B	B	B	B	A	B	A	A	B	B	B	B	A	B	A	A	B	B	A	A	B	B	B	B	B	B
98	B	A	B	B	B	B	B	A	B	B	B	A	B	A	B	A	B	B	B	B	B	B	A	B	B	B	B	A	B	B	B	B	B	B
99	A	B	B	B	B	B	B	B	A	B	B	A	A	B	A	B	B	B	B	A	A	A	A	B	B	B	A	A	B	B	A	B	B	B
100	A	B	B	B	B	B	B	A	B	A	B	A	A	B	A	B	B	B	B	A	A	A	A	A	B	B	B	B	B	B	A	B	B	B

7.3 Accordance to experimentation learning

m=1

player	FSA	FA	FSArate	FSB	FB	FSBrate	FS	F	FSrate	ACA	AA	ACArate	ACB	AB	ACBrate	AC	A	ACrate
1	5	27	0,19	7	32	0,22	12	59	0,20	16	21	0,76	13	19	0,68	29	40	0,73
2	9	19	0,47	20	32	0,63	29	51	0,57	12	21	0,57	9	27	0,33	21	48	0,44
7	8	8	1,00	17	27	0,63	25	35	0,71	15	34	0,44	5	30	0,17	20	64	0,31
8	14	30	0,47	0	27	0,00	14	57	0,25	15	34	0,44	4	8	0,50	19	42	0,45
9	10	23	0,43	11	24	0,46	21	47	0,45	12	25	0,48	12	27	0,44	24	52	0,46
10	16	27	0,59	13	24	0,54	29	51	0,57	16	25	0,64	16	23	0,70	32	48	0,67
11	17	25	0,68	18	25	0,72	35	50	0,70	10	23	0,43	11	26	0,42	21	49	0,43
12	17	26	0,65	14	25	0,56	31	51	0,61	11	23	0,48	10	25	0,40	21	48	0,44
13	23	27	0,85	0	4	0,00	23	31	0,74	13	45	0,29	13	23	0,57	26	68	0,38
14	22	23	0,96	2	4	0,50	24	27	0,89	8	45	0,18	6	27	0,22	14	72	0,19
15	19	25	0,76	13	17	0,76	32	42	0,76	13	30	0,43	16	27	0,59	29	57	0,51
16	16	27	0,59	12	17	0,71	28	44	0,64	10	30	0,33	15	25	0,60	25	55	0,45
21	18	30	0,60	13	21	0,62	31	51	0,61	6	26	0,23	10	22	0,45	16	48	0,33
22	7	22	0,32	5	21	0,24	12	43	0,28	14	26	0,54	13	30	0,43	27	56	0,48
23	49	49	1,00	0	0		49	49	1,00	2	46	0,04	2	4	0,50	4	50	0,08
24	4	4	1,00	0	0		4	4	1,00	16	46	0,35	16	49	0,33	32	95	0,34
27	8	18	0,44	5	13	0,38	13	31	0,42	10	29	0,34	12	39	0,31	22	68	0,32
28	36	39	0,92	9	13	0,69	45	52	0,87	12	29	0,41	10	18	0,56	22	47	0,47
29	18	29	0,62	26	37	0,70	44	66	0,67	6	19	0,32	6	14	0,43	12	33	0,36
30	5	14	0,36	27	37	0,73	32	51	0,63	12	19	0,63	10	29	0,34	22	48	0,46
mean	16,05	24,60	0,65	10,60	20,00	0,47	26,65	44,60	0,63	11,45	29,80	0,42	10,45	24,60	0,45	21,90	54,40	0,42

m=2

player	FSA	FA	FSArate	FSB	FB	FSBrate	FS	F	FSrate	ACA	AA	ACArate	ACB	AB	ACBrate	AC	A	ACrate
1	1	14	0,07	4	20	0,20	5	34	0,15	25	34	0,74	22	31	0,71	47	65	0,72
2	5	9	0,56	7	12	0,58	12	21	0,57	18	31	0,58	16	47	0,34	34	78	0,44
7	4	4	1,00	9	10	0,90	13	14	0,93	15	38	0,39	14	47	0,30	29	85	0,34
8	9	16	0,56	0	26	0,00	9	42	0,21	24	48	0,50	5	9	0,56	29	57	0,51
9	4	8	0,50	5	12	0,42	9	20	0,45	21	40	0,53	18	39	0,46	39	79	0,49
10	8	15	0,53	5	10	0,50	13	25	0,52	20	37	0,54	22	37	0,59	42	74	0,57
11	9	15	0,60	10	13	0,77	19	28	0,68	12	33	0,36	15	38	0,39	27	71	0,38
12	10	15	0,67	5	13	0,38	15	28	0,54	15	34	0,44	13	37	0,35	28	71	0,39
13	18	20	0,90	0	0		18	20	0,90	15	52	0,29	17	27	0,63	32	79	0,41
14	10	10	1,00	1	1	1,00	11	11	1,00	9	58	0,16	8	30	0,27	17	88	0,19
15	8	13	0,62	5	8	0,63	13	21	0,62	14	42	0,33	17	36	0,47	31	78	0,40
16	9	12	0,75	7	9	0,78	16	21	0,76	18	45	0,40	18	33	0,55	36	78	0,46
21	10	17	0,59	4	8	0,50	14	25	0,56	11	39	0,28	14	35	0,40	25	74	0,34
22	4	12	0,33	1	8	0,13	5	20	0,25	21	36	0,58	22	43	0,51	43	79	0,54
23	33	33	1,00	0	0		33	33	1,00	2	62	0,03	2	4	0,50	4	66	0,06
24	2	2	1,00	0	0		2	2	1,00	16	48	0,33	16	49	0,33	32	97	0,33
27	2	7	0,29	2	7	0,29	4	14	0,29	15	40	0,38	15	45	0,33	30	85	0,35
28	24	26	0,92	2	5	0,40	26	31	0,84	13	42	0,31	11	26	0,42	24	68	0,35
29	10	16	0,63	18	26	0,69	28	42	0,67	11	32	0,34	9	25	0,36	20	57	0,35
30	2	5	0,40	15	22	0,68	17	27	0,63	18	28	0,64	13	44	0,30	31	72	0,43
mean	9,10	13,45	0,65	5,00	10,50	0,52	14,10	23,95	0,63	15,65	40,95	0,41	14,35	34,10	0,44	30,00	75,05	0,40

m=3

player	FSA	FA	FSArate	FSB	FB	FSBrate	FS	F	FSrate	ACA	AA	ACArate	ACB	AB	ACBrate	AC	A	ACrate
1	0	6	0,00	0	12	0,00	0	18	0,00	32	42	0,76	26	39	0,67	58	81	0,72
2	2	2	1,00	1	4	0,25	3	6	0,50	22	38	0,58	18	55	0,33	40	93	0,43
7	0	0		1	1	1,00	1	1	1,00	15	42	0,36	15	56	0,27	30	98	0,31
8	6	8	0,75	0	25	0,00	6	33	0,18	29	56	0,52	6	10	0,60	35	66	0,53
9	0	2	0,00	3	4	0,75	3	6	0,50	23	46	0,50	24	47	0,51	47	93	0,51
10	3	7	0,43	1	2	0,50	4	9	0,44	23	45	0,51	26	45	0,58	49	90	0,54
11	4	8	0,50	3	4	0,75	7	12	0,58	14	40	0,35	17	47	0,36	31	87	0,36
12	7	8	0,88	2	6	0,33	9	14	0,64	19	41	0,46	17	44	0,39	36	85	0,42
13	13	14	0,93	0	0		13	14	0,93	16	58	0,28	17	27	0,63	33	85	0,39
14	5	5	1,00	0	0		5	5	1,00	9	63	0,14	8	31	0,26	17	94	0,18
15	3	6	0,50	0	3	0,00	3	9	0,33	16	49	0,33	17	41	0,41	33	90	0,37
16	2	4	0,50	5	5	1,00	7	9	0,78	19	53	0,36	20	37	0,54	39	90	0,43
21	7	10	0,70	2	3	0,67	9	13	0,69	15	46	0,33	17	40	0,43	32	86	0,37
22	1	5	0,20	0	1	0,00	1	6	0,17	25	43	0,58	28	50	0,56	53	93	0,57
23	21	21	1,00	0	0		21	21	1,00	2	74	0,03	2	4	0,50	4	78	0,05
24	0	0		0	0		0	0		16	50	0,32	16	49	0,33	32	99	0,32
27	0	3	0,00	0	2	0,00	0	5	0,00	17	44	0,39	18	50	0,36	35	94	0,37
28	14	16	0,88	2	4	0,50	16	20	0,80	13	52	0,25	12	27	0,44	25	79	0,32
29	5	8	0,63	13	18	0,72	18	26	0,69	14	40	0,35	12	33	0,36	26	73	0,36
30	1	1	1,00	7	11	0,64	8	12	0,67	21	32	0,66	16	55	0,29	37	87	0,43
mean	4,70	6,70	0,60	2,00	5,25	0,44	6,70	11,95	0,57	18,00	47,70	0,40	16,60	39,35	0,44	34,60	87,05	0,40

m=4

player	FSA	FA	FSArate	FSB	FB	FSBrate	FS	F	FSrate	ACA	AA	ACArate	ACB	AB	ACBrate	AC	A	ACrate
1	0	3	0,00	0	9	0,00	0	12	0,00	35	45	0,78	29	42	0,69	64	87	0,74
2	0	0		0	2	0,00	0	2	0,00	22	40	0,55	19	57	0,33	41	97	0,42
7	0	0		0	0		0	0		15	42	0,36	15	57	0,26	30	99	0,30
8	4	4	1,00	0	24	0,00	4	28	0,14	31	60	0,52	7	11	0,64	38	71	0,54
9	0	0		1	2	0,50	1	2	0,50	25	48	0,52	24	49	0,49	49	97	0,51
10	1	3	0,33	0	0		1	3	0,33	25	49	0,51	27	47	0,57	52	96	0,54
11	2	5	0,40	0	0		2	5	0,40	15	43	0,35	18	51	0,35	33	94	0,35
12	2	2	1,00	1	2	0,50	3	4	0,75	20	47	0,43	20	48	0,42	40	95	0,42
13	9	10	0,90	0	0		9	10	0,90	16	62	0,26	17	27	0,63	33	89	0,37
14	2	2	1,00	0	0		2	2	1,00	9	66	0,14	8	31	0,26	17	97	0,18
15	2	2	1,00	0	0		2	2	1,00	19	53	0,36	20	44	0,45	39	97	0,40
16	0	0		2	2	1,00	2	2	1,00	21	57	0,37	20	40	0,50	41	97	0,42
21	3	5	0,60	0	1	0,00	3	6	0,50	16	51	0,31	17	42	0,40	33	93	0,35
22	0	2	0,00	0	0		0	2	0,00	27	46	0,59	29	51	0,57	56	97	0,58
23	13	13	1,00	0	0		13	13	1,00	2	82	0,02	2	4	0,50	4	86	0,05
24	0	0		0	0		0	0		16	50	0,32	16	49	0,33	32	99	0,32
27	0	1	0,00	0	1	0,00	0	2	0,00	19	46	0,41	19	51	0,37	38	97	0,39
28	10	11	0,91	1	3	0,33	11	14	0,79	14	57	0,25	12	28	0,43	26	85	0,31
29	3	6	0,50	10	14	0,71	13	20	0,65	14	42	0,33	13	37	0,35	27	79	0,34
30	0	0		3	4	0,75	3	4	0,75	21	33	0,64	19	62	0,31	40	95	0,42
mean	2,55	3,45	0,62	0,90	3,20	0,29	3,45	6,65	0,54	19,10	50,95	0,40	17,55	41,40	0,44	36,65	92,35	0,40

F denotes “full record”
A denotes “ambiguous record”
S denotes “stay”
C denotes “change”
A, B denote strategy A and strategy B, respectively