Information and Beliefs in a Repeated Normal-form Game*

David Danz Dietmar Fehr

Social Science Research Center Berlin (WZB)

Dorothea Kübler

Social Science Research Center Berlin (WZB) & Technical University Berlin

forthcoming Experimental Economics

Abstract

We study beliefs and actions in a repeated normal-form game. Using a level-k model of limited strategic reasoning and allowing for other-regarding preferences, we classify action and belief choices with regard to their strategic sophistication and study their development over time. In addition to a baseline treatment with common knowledge of the game structure, feedback about actions in the previous period and random matching, we run treatments (i) with fixed matching, (ii) without information about the other player's payoffs, and (iii) without feedback about previous play. In all treatments with feedback, we observe more strategic play (increasing by 15 percent) and higher-level beliefs (increasing by 18 percent) over time. Without feedback, neither beliefs nor actions reach significantly higher levels of reasoning (with increases of 2 percentage points for actions and 6 percentage points for beliefs). The levels of reasoning reflected in actions and beliefs are highly consistent, but less so for types with lower levels of reasoning.

Keywords: experiments, beliefs, level-k model, learning

JEL classification numbers: C72, C92, D84

*We are grateful for valuable comments of two anonymous referees, as well as the editor, Jacob Goeree. We also thank Britt Grosskopf, Kyle Hyndman, Rajiv Sarin, Harald Uhlig, Roberto Weber, Georg Weizsäcker, Axel Werwatz and seminar participants at the Humboldt-Universität zu Berlin, European University Institute Florence, SFB 649 Workshop 2007, ESA World Meeting 2007, IMEBE 2008, VfS Annual Meeting 2008 and Econometric Society Meetings 2009 for valuable comments. We are indebted to Jana Stöver and Susanne Thiel for research assistence. Financial support from the Deutsche Forschungsgemeinschaft (DFG) through SFB 649 "Economic Risk" is gratefully acknowledged. *Corresponding Author:* Dorothea Kübler, Social Science Research Center Berlin (WZB), Reichpietschufer 50, 10785 Berlin, Germany. Email: *kuebler@wzb.eu.*

1 Introduction

The literature on learning in games has opened the black box of how an equilibrium is reached. Most learning models are backward looking and model decisions using past observations. More sophisticated learning models posit a deductive reasoning process implying that players analyze the game in order to understand its strategic properties and thereby form beliefs about the other player's choice. This paper studies such learning processes and any combinations of them by tracking both action choices and beliefs in a repeated normal-form game. We ask whether actions and beliefs develop in a consistent manner under various treatment conditions.

The game we implemented in our experiments allows for a distinction between actions and beliefs according to their levels of reasoning in the sense of the level-k model (Nagel, 1995; Stahl and Wilson, 1995). The level-k model assumes that players differ in their beliefs about other players' strategic sophistication. The existing literature analyzes levels of reasoning in one-shot games under the assumption of self-interested preferences. We depart from this literature in two ways. First, since the classification of actions as level-1, level-2 etc. as well as the assessment of best response behavior relies on a player's preferences, it is necessary to take possible preference heterogeneity into account to avoid possible misspecifications. We propose an approach to study levels of reasoning while allowing for other-regarding preferences.[1] Second, we use the level-k model to track *changes* in the level of strategic behavior over time and to check whether actions and beliefs develop in a consistent manner. To understand how learning takes place, i.e., whether higher levels of reasoning are mainly reached through introspection or through observation of previous play or both, we vary the information conditions.[2]

In our experiment, participants have full information about the normal-form game and receive feedback about their own payoff (and thereby the other's payoff and action) in the previous period. In order to separate between different sources of learning, we employ control treatments. In one control treatment, subjects receive no feedback about previous play (see e.g. Weber 2003, Rick and Weber 2010).[3] In another control treatment, subjects receive feedback about previous

[1]If subjects have other-regarding or efficiency-oriented preferences, the game we study has multiple equilibria. However, the level-k model is well suited to make a unique prediction for such games, thereby avoiding a selection problem.

[2]The level-k model has been associated with introspection by Goeree and Holt (2004). For one-shot games, introspection or deductive reasoning is clearly the obvious choice. However, when levels of reasoning are analyzed in repeated interactions, feedback may affect beliefs and may thereby lead to higher-level play.

[3]It is conceivable that experience and observation of past play reduce sophistication relative to play without feedback. In a feedback-free environment, subjects are presumably forced to think about the game and therefore they may acquire simple solution concepts such as iterated dominance or backwards induction. Rick and Weber (2010)

play, but do not know the payoff function of the other player.[4] The two information conditions thus allow us to compare the relative importance of the deductive and inductive elements of learning to play strategically within the normal-form game we have chosen.

We observe an initially high level of nonstrategic actions in all treatments, implying that subjects tend to neglect the incentives of the other player. However, in the three treatments with feedback about the other player's past behavior, the proportion of nonstrategic actions decreases over time (on average by 15 percentage points). Similarly, the proportion of subjects holding non-strategic level-1 beliefs decreases on average by 18 percentage points in the treatments with feedback, a result that is confirmed when estimating beliefs as noisy responses to the true underlying beliefs (as in Costa-Gomes and Weizsäcker, 2008). Without feedback about past outcomes, strategic play increases only by 2 percentage points and strategic beliefs by 6 percentage points. Thus, both the action and belief data show that feedback is more important for the learning process than information about the payoffs of the opponent in the game we are studying. For most level-k types and treatments, beliefs and actions are consistent with each other over time in the sense that subjects choose actions according to their belief type as their level of reasoning increases. However, this consistency is lower for types with lower levels of reasoning across all treatments.

There are a few recent papers using belief elicitation procedures in finitely repeated normal-form games. Nyarko and Schotter (2002) investigated the explanatory power of beliefs inferred from belief-learning models. They used a 2×2 game with a unique mixed-strategy equilibrium and found that stated beliefs describe the learning path better than beliefs constructed from previous actions with belief-learning models. Moreover, this difference between stated beliefs and constructed beliefs does not decrease over time. The two other papers closest to our design are Hyndman et al. (forthcoming) and Terracol and Vaksmann (2009). In contrast to our experiment, these two papers use games where the Nash equilibrium is on the Pareto frontier and focus on strategic teaching and its underlying mechanisms. They demonstrate that a sophisticated subject can act as a teacher and thereby facilitate the convergence process to an equilibrium in the presence of a sufficiently fast follower. All three studies do not attempt to analyze the co-development of actions and beliefs over time in terms of levels of reasoning. In addition, they do not vary the available information

demonstrate that subjects are able to acquire and to transfer such concepts to similar games.

[4] Oechssler and Schipper (2006) used a similar setup to study subjects' ability to learn about the game payoffs. So-called "minimal social situations", introduced by Sidowski (1957), have been studied widely by psychologists. They differ from our control treatment in that players do not know their own payoffs, nor the other player's strategy set. Often they are not even aware that they are in a situation with interdependent payoffs.

Table 1: Normal-form game used in experiment.

	Left (L)	**Center (C)**	**Right (R)**
Top (T)	78, 68	72, 23	12, 20
Middle (M)	67, 52	59, 63	78, 49
Bottom (B)	21, 11	62, 89	89, 78

in a systematic way in order to identify the sources of learning.[5]

The paper is organized as follows. The next section introduces the design and procedures of the experiment and provides a description of the level-k model applied to the normal-form game we used. In Section 3, we present the results, focusing first on actions, then on belief statements and finally on the consistency of actions and beliefs. Section 4 concludes.

2 Experimental design

2.1 The game and strategies

In the experiment, we used the asymmetric normal-form game presented in Table 1. The game has a unique Nash equilibrium in pure strategies in which the row player chooses Top and the column player chooses Left. This Nash equilibrium of the stage game is not Pareto efficient. The strategy combination of Bottom and Right not only maximizes the payoff of the player who is least well off, but also maximizes the sum of payoffs. However, for the column player choosing Right is strictly dominated by Center.[6] Since the joint payoff maximizing and maxmin outcome (Bottom, Right) cannot be accounted for with standard preferences, we allow for other-regarding preferences by using a simple linear specification in the spirit of Edgeworth (1881)[7]:

$$u_i(\pi_i, \pi_j) = \pi_i + \theta\pi_j. \tag{1}$$

Here, π_i denotes the player's own payoff and π_j denotes the payoff of the other player, and $\theta \in [0, 1]$ is the willingness to exchange own for other's payoff. Besides risk neutrality we assume

[5] Hyndman et al. (forthcoming) ran a treatment with incomplete information about the other player's payoffs, but none of the mentioned papers employs a treatment without feedback.

[6] The unique Nash equilibrium of the stage game is also the unique subgame perfect equilibrium of the repeated game. However, there exist Nash equilibria of the finitely repeated game with fixed matching that are not subgame perfect. For example, playing the Pareto-efficient strategy combination (Bottom, Right) for at most 17 periods and then reverting to the Nash equilibrium (Top, Left) for the remaining three periods is a Nash Equilibrium.

[7] We are grateful to a referee for pointing out that this widely-used model is originally due to Edgeworth. For more recent applications see Anderson et al (1998), Goeree et al (2002) or Cox et al (2007).

Table 2: Predicted actions and beliefs by strategic sophistication and social preferences.

		Actions			Beliefs			
	θ–interval	$L1$	$L2$	$L3+$	$L1$	$L2$	$L3$	$L4+$
Row player	$[0.00, 0.15)$	M	T	T	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 1, 0)$	$(1, 0, 0)$
	$[0.15, 0.40)$	M	B	B	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
	$[0.40, 0.73)$	M	B	B	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 1, 0)$	$(0, 0, 1)$
	$[0.73, 1.00]$	M	B	B	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 0, 1)$	$(0, 0, 1)$
Column player	$[0.00, 0.15)$	C	C	L	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
	$[0.15, 0.40)$	C	C	C	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 0, 1)$	$(0, 0, 1)$
	$[0.40, 0.73)$	C	C	R	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 0, 1)$	$(0, 0, 1)$
	$[0.73, 1.00]$	C	R	R	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 0, 1)$	$(0, 0, 1)$

Notes: Row player's action are T(op), M(iddle), B(ottom) and column player's action are L(eft), C(enter), R(ight). Beliefs are denoted as vectors, for example, (0,1,0) is the degenerate belief that the other player chooses M and C, respectively.

that players whose utility function is characterized by a certain θ expect the other player to have the same utility function and value of θ.

The main focus of this study is on the development of strategic and nonstrategic behavior over several periods of play. The level-k model provides a parsimonious classification of actions and beliefs in terms of their strategic sophistication. It assumes that players differ in their level of reasoning, which is modeled as an iterative best-response process. The lowest level is a non-strategic player (level-0) who randomizes uniformly over his strategy space and a higher-level player best responds to a level-$(k-1)$ type with $k \in \{1, 2, .., \infty\}$.[8] Given the utility formulation (1), it is straightforward to derive the predictions for actions and belief statements of the level-k model (see Table 2).

The table shows that the predictions depend on the parameter θ. Since in our dynamic setting, a player's parameter θ may change in the course of the experiment as a reaction to the history of play, we have to be careful not to confound such changes in θ with learning. For example, a column player choosing L(eft) has a level of reasoning of at least 3 with $\theta < 0.15$, denoted by $L3^+_{\theta\leq .15}$ (see Table 3). Hence, if we observe an increase of $L3^+_{\theta\leq .15}$ play over time for column players, this could be due to (i) subjects reaching higher levels of reasoning if θ is time-invariant or (ii) a decrease in θ which might not be accompanied by learning at all, or a combination of both.

Nevertheless, we are able to infer the level of reasoning of actions and beliefs independently

[8] While in Stahl and Wilson's formulation a level-k type best responds to a distribution of lower level types, we use the described alternative formulation, as for example, in Nagel (1995) or Costa-Gomes et al (2001). The level-k model has been tested and extended by various other studies in the context of one-shot normal-form games (see, among others, Costa-Gomes et al 2001; Costa-Gomes and Weizsäcker 2008; Rey Biel 2009; and Camerer et al 2004). The most common types found in normal-form games are level-1, level-2 and Nash types.

Table 3: Identification of player types by means of actions and beliefs.

	Actions			Beliefs			
	M/C	B/R	T/L	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(0,1,0)$	$(0,0,1)$	$(1,0,0)$
Row player	$L1$	$L2^{+}_{\theta>.15}$	$L2^{+}_{\theta\leq.15}$	$L1$	$L2^{+}$	$L3^{+}_{\theta>.40}$	$L4^{+}_{\theta\leq.15}$
Column player	$L1^{+}$	$L2^{+}_{\theta>.40}$	$L3^{+}_{\theta\leq.15}$	$L1$	$L2$	$L3^{+}_{\theta>.15}$	$L3^{+}_{\theta\leq.15}$

of θ for a subset of choices. First, row players of type $L1$ will exclusively choose action M for any value of $\theta \in [0, 1]$. Similarly, denoting a belief over the other players choice set {L,C,R}, and {T,M,B} respectively, as $b = (b_1, b_2, b_3) \in \Delta^2$, stated beliefs of $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$ for either player role are solely predicted for $L1$ types for all $\theta \in [0, 1]$. Furthermore, a column player's belief statement of $(0, 1, 0)$ is predicted for $L2$ players only, again for all $\theta \in [0, 1]$. Thus, regarding the action choices, the row player provides us with clear evidence on learning, as a decrease in $L1$ play cannot be due to changes in θ. On the other hand, regarding the beliefs, learning that is not confounded with changes in θ is reflected in changes in $L1$ as well as $L2$ for the column player. For the row player only changes in $L1$ beliefs are not confounded with potential changes of θ over time. Regarding the remaining action and belief types listed in Table 3, we will be cautious with the interpretation of their development over time as learning, since it might be confounded with changes in θ.

2.2 Treatments and Procedures

The level-k model allows us to measure learning in the sense of subjects stating higher-level beliefs and choosing higher-level strategies over time. Suppose for example that a subject starts out by playing the level-1 ($L1$) action. In our repeated setting, she may use the available information about the previous actions or the incentives of the other player to reason about future play. It can therefore be expected that she learns to best respond to $L1$ by playing $L2$ and so forth, depending on the amount of information available.

We implemented four treatments that differ with respect to the information available for the participants (for a detailed overview of the treatments see Table 4). Our baseline is the random-matching treatment (RM), where subjects had full information about the game. In addition, after each period they received feedback about the payoff earned in this period (and thereby about the action of the other player). In the second treatment, we only changed the matching scheme to fixed matching, denoted by FM, in order to understand whether interaction with the same partner speeds up the learning process. The two remaining treatments both use a fixed-matching protocol.

Table 4: Overview of treatments.

Treatment	Payoff	Feedback	Matching	Periods	Sessions	# of subjects
RM	own&other player	own payoff	random	20	4	54
FM	own&other player	own payoff	fixed	20	4	54
PI	own	own payoff	fixed	20	4	48
NF	own&other player	none	fixed	20	4	50

In treatment NF (no feedback), subjects received no feedback at all, but had common knowledge of the payoff structure of the game as in RM and FM. Since there is no feedback in NF, the matching protocol does not matter for the game-theoretic prediction. In treatment PI (partial information), subjects were only informed about their own payoff function. They could, however, infer the choice of the other player through the feedback received after each period, but not the other players' payoff. Thus treatment PI provides an environment where only the observation of past actions can inform the learning process.[9] Note that repeated-game effects are in principle possible in treatment FM, but not in RM, NF and PI. Without feedback in NF or without information about the payoffs of the other player in PI, strategies that punish a player for deviations from the equilibrium path are impossible.

Subjects were recruited with ORSEE (Greiner, 2004) from the student population at Technical University Berlin. The experiment was computerized using z-Tree (Fischbacher, 2007). Sessions lasted about one hour, and subjects' average earnings were about € 12.80. At the beginning of a session, subjects were randomly assigned a fixed player role (row or column).[10] In each period, subjects had to make two decisions. First, they had to indicate their beliefs regarding the behavior of the other player. Specifically, we asked subjects to state the expected frequencies of play, i.e., they had to specify in how many out of 100 times they expected the other player to choose the three actions in the current period. After the belief task, subjects had to make their choice by selecting one of the three possible actions.

For the action choice, subjects were paid according to the numbers in the payoff matrix (Table 1), which were exchanged at the commonly known rate of 1 point = € 0.15, e.g., the Nash equilibrium (Top, Left) would lead to payoffs of € 11.70 and € 10.20. The maximum payoff for the

[9] We use the labels $L1$, $L2^{+}_{\theta<.15}$ etc. also in treatment PI even though a priori the subjects are not aware of the other player's payoffs and other-regarding preferences are less relevant. However, subjects can use their feedback to construct a "subjective game". Kalai and Lehrer (1993) show that subjective games can converge to an ε-Nash equilibrium of the underlying game.

[10] Upon entering the lab, subjects received written instructions including an understanding test. The experiment only started after all subjects had answered the questions correctly. For a sample of the instructions see the Appendix.

belief task was € 1.50 and we used a standard quadratic scoring rule, which is incentive compatible given that subjects are risk-neutral money maximizers. Subjects did not receive any feedback about their payoffs from the belief elicitation task.[11] Paying subjects for the choice and the belief task may lead them into following a combined strategy for both tasks together over all periods. To discourage such behavior, we randomly select one payoff-relevant period for the action task and one payoff-relevant period for the belief task at the end of the experiment.[12]

3 Results

3.1 Descriptive statistics of actions

Figure 1 presents the actions aggregated over all periods for each treatment and player role.[13] Assuming that row and column players are drawn from the same population, the implications of the level-k model regarding the relative frequencies of actions across player roles can be tested (see Table 2). In all four treatments, the proportion of $L1^+$ actions (Center) of column players is higher than the fraction of $L1$ actions (Middle) of row players, as expected. Moreover, in all treatments, there are more $L2^+_{\theta>.15}$ actions (Bottom) of the row player than $L2^+_{\theta>.40}$ actions (Right) of the column player, consistent with the size of the interval of θ. Finally, we observe more $L2^+_{\theta\leq.15}$ play (Top) by row players than $L3^+_{\theta\leq.15}$ play (Left) by column players, again as expected when row and column players are drawn from the same distribution of level-k types.

Regarding treatment effects, two observations seem noteworthy. First, in treatments FM and RM the row players' behavior is statistically indistinguishable ($\chi^2_{(2)}$, $p = 0.13$), whereas the column players' behavior is affected by the matching protocol resulting in more $L1^+$ play in RM than in FM.[14] This difference can be ascribed to the fact that the column player's action Right is more risky in the stranger design of RM than in FM with a fixed partner where column players

[11] Of course, subjects could infer their belief-elicitation payoff from the game feedback. The main reason for not providing the belief-elicitation payoff was to change as few parameters as possible when going from RM, PI and FM to NF.

[12] Subjects may for example coordinate on some non-equilibrium outcome and become unwilling to move away from it in order to avoid losses in the belief task (lock-in effect) or they may try to hedge with the stated belief against adverse outcomes of the action task within a given period. Blanco et al (2010) investigate the potential of hedging and find no evidence for it unless the hedging opportunity is very salient. Their hedging-proof procedure either pays the decision or the belief-elicitation task randomly.

[13] When considering first-period choices, we can compare our results to experiments using one-shot games. Pooling the data over player roles, we observe 51% $L1/L1^+$ behavior in the first period in RM, FM and NF. This is in line with previous studies. For instance, Costa-Gomes et al. (2001) observed a frequency of $L1$ actions of about 45%, Rey-Biel (2009) found 48% $L1$ behavior in constant-sum games, whereas Costa-Gomes and Weizsäcker (2008) observed slightly higher rates of about 60%.

[14] All results reported as significant in the paper are based on a 5%-level of significance.

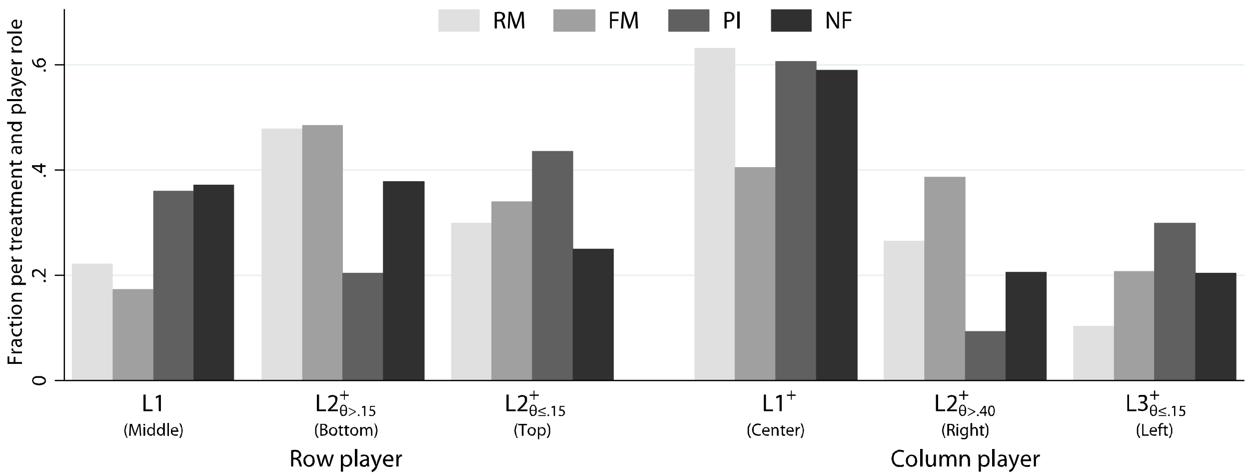


Figure 1: Actions aggregated over all periods.

"invest" in the cooperative outcome. Second, we observe that in treatment PI the proportion of nonstrategic behavior ($L1/L1^+$) is higher than in FM. Moreover, in PI there is more $L2^+_{(\theta\leq.15)}$ than $L2^+_{(\theta>.15)}$ play by the row players and more $L3^+_{(\theta\leq.15)}$ than $L2^+_{(\theta>.40)}$ play by the column players while the reverse is true in all other treatments. The higher proportion of non-strategic and the lower proportion of other-regarding choices in PI can be explained by the lack of information about the other player's payoffs.

3.2 Actions over time

Figure 2 presents the development of actions over time for each treatment. The figure shows averages over three periods in a given treatment for row players in the top panels and for column players in the bottom panels. For the row player, we observe a decrease in nonstrategic play ($L1$) in most treatments (top-left panel). In both RM and FM the decline from the first three to the last three periods is 16 percentage points, and in PI nonstrategic play declines by 18 percentage points. Only in NF there is no change at all. The decrease of nonstrategic play in RM, FM, and PI is associated with an increase in strategic play with selfish preferences (top-right panel). In particular, $L2^+_{\theta\leq.15}$ increases in PI by 24 percentage points, in RM by 19 percentage points, in FM by 16 percentage points. There is almost no change in strategic behavior with non-selfish preferences ($L2^+_{\theta>.15}$). Again there is no change in NF. For the column player, learning to choose higher-level actions cannot be identified for intermediate values of $\theta\in[0.15, 0.40)$. Thus, without an assumption on the development of θ over time, the observed decrease in $L1^+$ actions in some treatments as

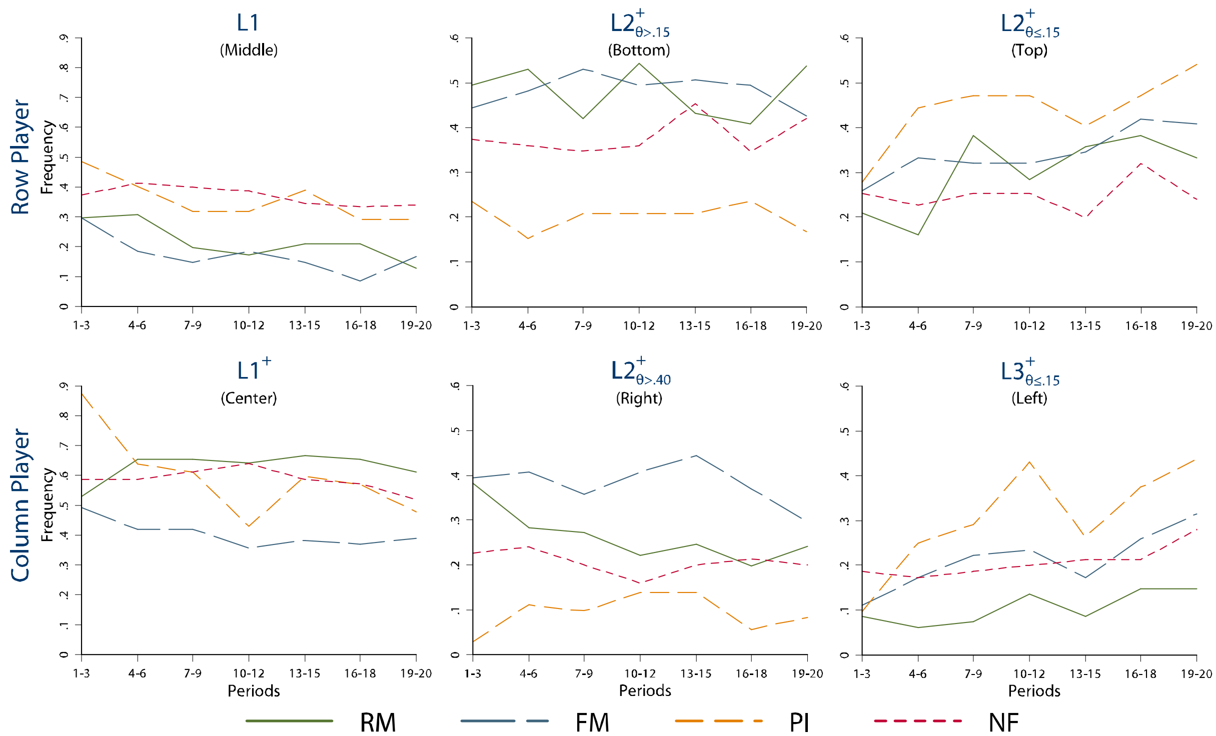


Figure 2: Levels of reasoning determined by actions over time.

well as the increase in strategic actions with selfish preferences $L3^{+}_{\theta\leq.15}$ in all treatments cannot unambiguously be ascribed to learning. This observed increase in $L3^{+}_{\theta\leq.15}$ actions is strongest in FM and PI (with 21 and 33 percentage points) while in RM and NF it increases only modestly by 8 percentage points each. Similar as for row players, we observe weaker changes in non-selfish behavior, $L2^{+}_{\theta>.40}$, although in RM and FM $L2^{+}_{\theta>.40}$ decreases by 17 and 6 percentage points, respectively.

Aggregating over both player roles and all treatments with feedback, we observe an increase in strategic play (i.e., a decrease in $L1/L1^{+}$ play) between the first and last three periods of 15 percentage points (for row players alone 17 percentage points) as opposed to the no-feedback treatment with a decrease in 2 percentage points. But keep in mind that these fractions might underestimate learning because we count $L1^{+}$ actions as nonstrategic.

For statistical evidence of the potential learning paths, we run probit regressions for each player role. The dependent variable is an indicator variable for each action. The regressions contain an intercept and a time trend for each treatment separately. We report the estimated time trends in Table 5. Again, we are mainly interested in row player behavior since a decrease in $L1$ play

Table 5: Time trends of levels of reasoning determined by actions.

	Row player			Column player		
	L1 (Middle)	$L2^+_{\theta>.15}$ (Bottom)	$L2^+_{\theta\leq.15}$ (Top)	$L1^+$ (Center)	$L2^+_{\theta>.40}$ (Right)	$L3^+_{\theta\leq.15}$ (Left)
RM*Period	-0.039^{***}	-0.011	0.040^{***}	0.014	-0.041^{***}	0.023^{*}
	(0.012)	(0.011)	(0.011)	(0.011)	(0.013)	(0.014)
FM*Period	-0.044^{***}	0.001	0.032^{***}	-0.019^{*}	-0.013	0.042^{***}
	(0.013)	(0.011)	(0.012)	(0.010)	(0.012)	(0.013)
PI*Period	-0.041^{***}	-0.002	0.033^{***}	-0.050^{***}	0.007	0.053^{***}
	(0.012)	(0.012)	(0.011)	(0.011)	(0.016)	(0.012)
NF*Period	-0.013	0.007	0.008	-0.008	-0.012	0.019
	(0.011)	(0.012)	(0.012)	(0.011)	(0.013)	(0.013)
N		2060			2060	
logL	-958.74	-1041.38	-1030.61	-1191.98	-837.84	-833.44

Notes: Probit regressions with individual random effects (standard errors in parentheses). Intercepts are omitted in the table. * p<0.10, ** p<0.05, *** p<0.01

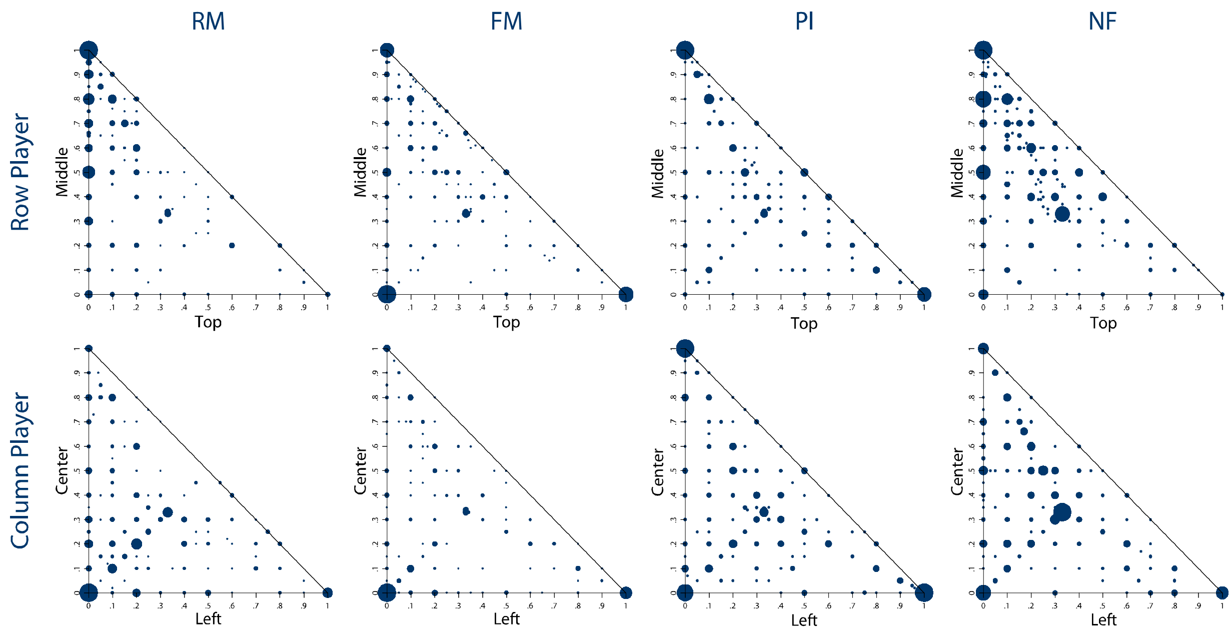
cannot be due to a change in θ. In all treatments with feedback (RM, FM and PI) we observe significant negative time trends in $L1$ actions.[15] A Wilcoxon signed-rank test to test for differences in the observed fraction of row players' $L1$ choices in the first five versus the last five periods yields the same significant decreases. For column players in PI, behavior should not be confounded with changes in θ because they lack information about row players' payoffs. Indeed, the regressions indicate a strong decrease in $L1^+$ and an increase in $L3^+_{\theta\leq.15}$ in PI. Similar trends are observed in treatment FM, and the non-parametric approach confirms an increase of $L3^+_{\theta\leq.15}$ play in FM and in PI.

Result 1 *Over time, we observe significant increases of strategic play in the treatments with feedback (RM, FM and PI). Without feedback (treatment NF), there is only a slight but insignificant trend towards higher-level play.*

3.3 Descriptive statistics of stated beliefs

Figure 3 shows a two-dimensional simplex of beliefs for each treatment splitted by player role, in which each vertex denotes a degenerate belief. The level-k model predicts that subjects' beliefs about the other player's actions are taken from the belief set $B = \{(1/3, 1/3, 1/3), (0, 1, 0), (0, 0, 1), (1, 0, 0)\}$,

[15] Note that at first sight our results in treatment NF differ somewhat from the results of Rick and Weber (2010) who find significant learning without feedback information. However, in their repeated asymmetric 3×3 game which is closest to the game we use, the absolute changes in action choices are comparable in size to our results.



Note: Each vertex denotes a degenerate belief, e.g., the point (0,0) represents the belief (0,0,1) that the column (row) player will play B (R).

Figure 3: Stated beliefs by treatment and player role.

see Table 2. Figure 3 reveals that the modes of the subjects' stated beliefs are indeed equal to these predicted beliefs. In particular, the stated beliefs exactly coincide with the predictions in 37% of all cases, and many of the remaining belief statements are in the proximity.

In order to classify a stated belief, we calculate its Euclidean distance to all four level-k beliefs. We then assign the stated belief to the level-k belief with the smallest distance. Figure 4 depicts the frequencies of the categorized belief statements by treatment and player role.

The level-k model has implications regarding the relative frequencies of belief types across player roles if we again assume that row and column players come from the same distribution of level-k types. Specifically, we observe that the fractions of $L1$ beliefs are not different between row and column players in any treatment.[16] Moreover, there is a larger fraction of row players' $L2^+$ beliefs than column players' $L2$ beliefs in every treatment, which is consistent with the larger $L2^+$-belief set. Also, as expected due to the smaller interval of θ, row players with $L3^+_{\theta>.40}$ beliefs are observed less often than column players with $L3^+_{\theta>.15}$ beliefs. Finally, the fraction of row players with $L4^+_{\theta\leq.15}$ beliefs is lower than the fraction of column players with $L3^+_{\theta\leq.15}$ beliefs in all treatments.[17] The observed consistency can be interpreted as supporting the descriptive validity

[16]We run probit regressions with $L1$-beliefs as the dependent variable and a dummy for player role as the independent variable where we corrected for repeated observations on the individual level. The lowest p-value is found in RM ($p = 0.167$).

[17]Similarly, aggregated comparisons between *actions* and *beliefs* (such as the row player's fractions of Middle

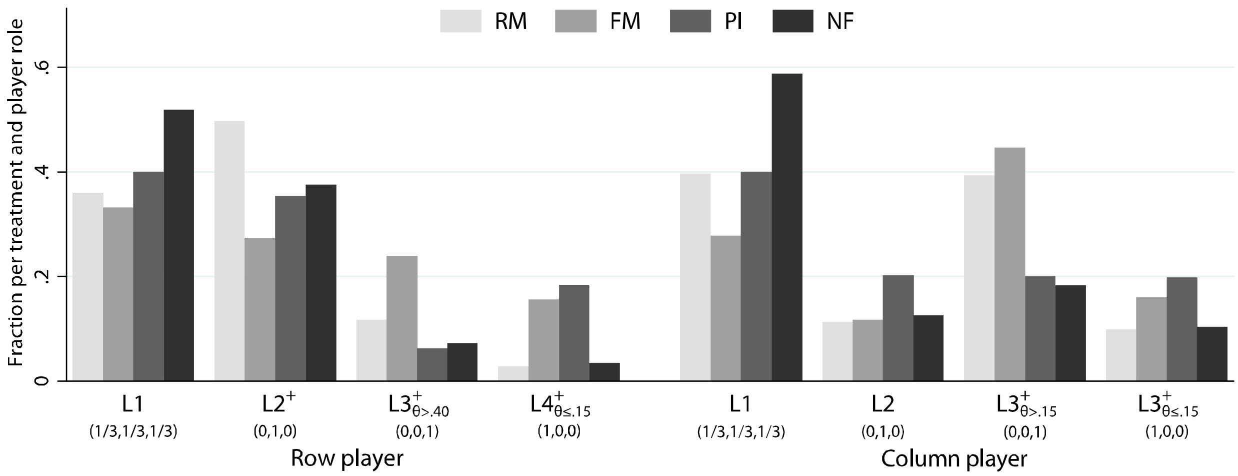


Figure 4: Levels of reasoning determined by stated beliefs over all periods.

of the level-k model with prefererence heterogeneity.

Apart from the overall high level of nonstrategic $L1$ beliefs, Figure 4 shows that both row and column players exhibit the highest fraction of nonstrategic beliefs in NF followed by PI, RM and FM (in this order). Thus, the lack of feedback (NF) leads to lower-level beliefs in terms of the level-k model.

3.4 Beliefs over time

Figure 5 depicts the development of strategic sophistication measured by stated beliefs over time. Similar to the action choices, we find a decrease of $L1$ beliefs for almost all treatments (left panels in Figure 5) as well as an increase in higher-level beliefs with selfish preferences (right panels in Figure 5).

For row players the decrease in $L1$ beliefs, measured as the difference between the first and the last three periods is 21 percentage points in RM and 16 percentage points in FM, and about 11 percentage points in both PI and NF. Remember that changes in $L1$ beliefs are evidence for learning, independent of θ, for both row and column players. For column players we observe a decline of $L1$ beliefs in PI of 31 percentage points and of 25 percentage points in FM. In RM and NF the decline is only modest with 6 and 1 percentage points, respectively. We observe a similar increase in strategic beliefs as the $L3^+_{\theta\leq.15}$ and $L3^+_{\theta>.15}$ beliefs of the column player increase over

choices and $L1$ beliefs) are consistent with a unique underlying distribution of levels of reasoning (see Figures 1 and 4). Together with the reported comparisons *within* the action as well as the belief space, only 4 out of 72 possible tests (5.6%) on the treatment level contradict the level-k predictions.

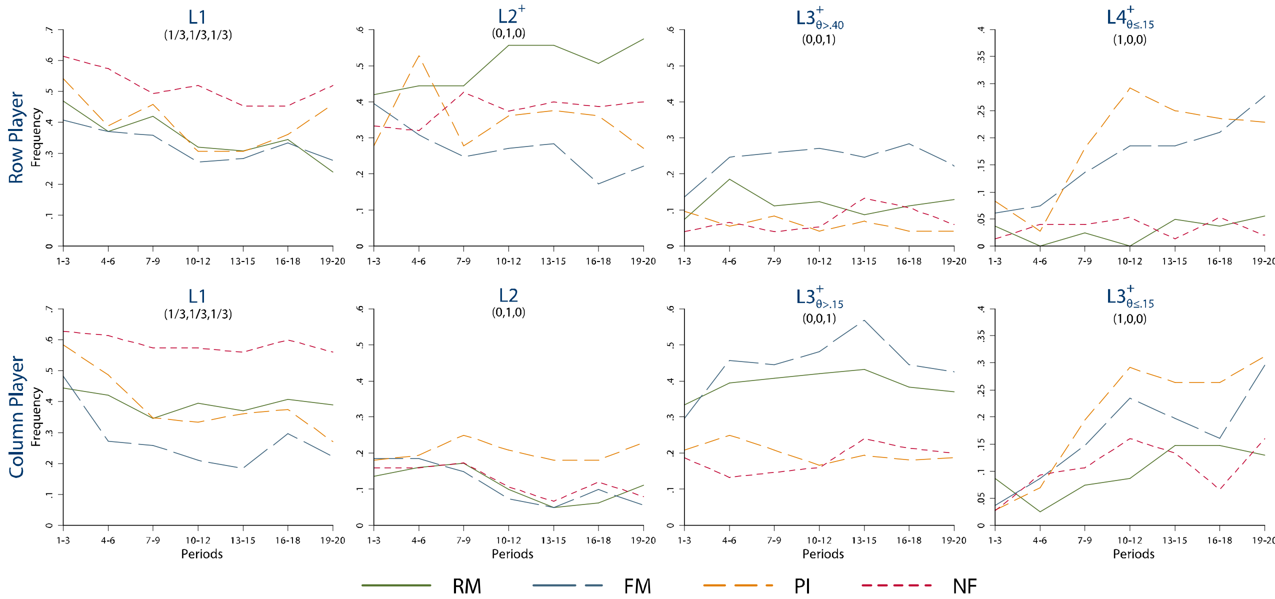


Figure 5: Levels of reasoning determined by stated beliefs over time.

time by 14 and 11 percentage points in RM, and in FM by 22 and 14 percentage points, which is again evidence for learning independent of θ. Finally, $L1$ beliefs are highest in NF and decrease only very little.

On average for both player roles in all treatments with feedback, we observe an increase in strategic beliefs and a decrease in $L1$ beliefs between the first and the last three periods of 18 percentage points as opposed to the no-feedback treatment with a decrease of merely 6 percentage points.

In order to analyze how the subjects' level of reasoning expressed by the belief statements changes over time, we employ a statistical model to capture possible errors in belief statements. This model has a number of advantages over the simple classification of beliefs according to their distance to the predicted belief types that we have employed before. Most importantly, the model weighs the evidence for a certain belief type differently depending on the precise distance of a stated belief from each predicted belief. In the model which we adopt from Costa-Gomes and Weizsäcker (2008), we assume that players hold an unobservable underlying belief $b^u = (b^u_1, b^u_2, b^u_3) \in \Delta^2$ about the other player's action. In accordance with the implemented quadratic scoring rule the expected payoff from a belief statement $b^s = (b^s_1, b^s_2, b^s_3) \in \Delta^2$ given the true underlying belief b^u is given by $E(\pi(b^s|b^u)) = c - d\sum_{i=1}^3 b^u_i(\sum_{j=1}^3(I_{i=j} - b^s_j)^2)$, which is maximized at $b^S = b^u$, i.e., truth-telling is optimal for risk-neutral players. However, we do not assume that subjects necessarily choose the exact optimum. Instead, we assume that the subjects give noisy best responses to their underlying beliefs following the logistic density

$$r(b^s|b^u,\lambda) = \frac{\exp(\lambda E(\pi(b^s|b^u)))}{\int_{s\in\Delta^2}\exp(\lambda E(\pi(s|b^u)))ds}, \tag{2}$$

where λ governs the noise in the best response. For $\lambda \to 0$ the player's belief statements follow a uniform distribution over the set of permitted belief statements whereas for $\lambda \to \infty$ the player's stated belief converges in probability to her underlying belief.

Following the predictions of the level-k model, we assume four types of underlying beliefs denoted by $b^u \in B = \{b_1^u, b_2^u, b_3^u, b_4^u\} = \{(1/3, 1/3, 1/3), (0, 1, 0), (0, 0, 1), (1, 0, 0)\}$, with a slight abuse of notation. Using a mixture model approach, we estimate the probability weights $w = (w_1, w_2, w_3, w_4)$ for each underlying belief type. Since we are interested in the development of these weights over time, we consider a log likelihood which includes a time trend $w^{\Delta t} = (w_1^{\Delta t}, w_2^{\Delta t}, w_3^{\Delta t}, w_4^{\Delta t})$ for each type weight together with the corresponding weight estimates for the first period $w^{t1} = (w_1^{t1}, w_2^{t1}, w_3^{t1}, w_4^{t1})$:

$$L(w^{t1}, w^{\Delta t}, \lambda|b^u, b^s) = \sum_{i=1}^{N}\sum_{t=1}^{20} \ln\left[\sum_{k=1}^{4}\left(w_k^{t1} + w_k^{\Delta t}(t-1)\right) r(\lambda|b_{it}^s, b_k^u)\right] \tag{3}$$

Maximizing (3) for each treatment and player role separately yields the trend estimates $\widehat{w}^{\Delta t}$ as listed in Table 6. All estimates of λ are significantly different from zero except for the row player in RM (ranging from 7.53 to 28.6; not reported in the table). In addition, all parameter estimates of the time trends are in line with learning to hold higher-level beliefs whenever the estimates are significant. Both findings can be seen as support for the model we have chosen.

The first row in Table 6 shows that movements toward higher-level beliefs are observed in all treatments. Our main focus is on column players to identify learning with respect to beliefs since changes in $L1$ as well as $L2$ are not confounded with changes in θ. There is a significant decrease of $L1$ and $L2$ beliefs in FM in favor of $L3^+_{\theta\leq.15}$ beliefs, and a significant decrease of $L1$ beliefs in PI again in favor of $L3^+_{\theta\leq.15}$ beliefs for column players.

Assuming that θ is constant over time, the row players reveal a significant increase of $L2^+$ beliefs in RM, which is mostly due to the marginal decrease in $L1$ beliefs. Moreover, similar to the column players, the row players show an increase of $L4^+_{\theta\leq.15}$ beliefs in FM and PI. Finally, in NF there is no significant change in the strategic sophistication of beliefs neither for row nor column

Table 6: Time trends of levels of reasoning determined by stated beliefs.

	Row player					Column player			
	RM	FM	PI	NF		RM	FM	PI	NF
L1	-0.008^{*}	-0.007^{*}	-0.010^{*}	-0.008^{*}	L1	-0.002	-0.010^{**}	-0.015^{***}	-0.005
$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$	(0.005)	(0.004)	(0.005)	(0.005)	$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$	(0.004)	(0.004)	(0.006)	(0.004)
$L2^{+}$	0.010^{**}	-0.009^{***}	-0.002	0.003	L2	-0.007	-0.008^{**}	-0.000	-0.004
(0,1,0)	(0.005)	(0.003)	(0.006)	(0.005)	(0,1,0)	(0.004)	(0.003)	(0.005)	(0.003)
$L3^{+}_{\theta>.40}$	-0.003	0.005	-0.003	0.004	$L3^{+}_{\theta>.15}$	0.003	0.005	-0.003	0.003
(0,0,1)	(0.004)	(0.005)	(0.002)	(0.003)	(0,0,1)	(0.007)	(0.005)	(0.006)	(0.004)
$L4^{+}_{\theta\leq.15}$	0.001	0.011^{***}	0.014^{**}	0.001	$L3^{+}_{\theta\leq.15}$	0.006	0.013^{**}	0.018^{***}	0.006
(1,0,0)	(0.002)	(0.004)	(0.006)	(0.002)	(1,0,0)	(0.004)	(0.005)	(0.006)	(0.004)
N	540	540	480	500		540	540	480	500
log L	701.84	730.93	603.97	530.35		636.71	913.59	648.61	458.31

Notes: The reported coefficients are estimates of the time trends in the type weights $w^{\Delta t}$. Standard errors in parentheses and corrected for observational clusters on the individual level. * p<0.10, ** p<0.05, *** p<0.01.

players. Wilcoxon signed-rank tests of the differences in the observed fraction of each belief type in the first five versus the last five periods overall corroborate these results.[18]

Result 2 *Over time, higher-level beliefs increase significantly in treatments with feedback and fixed partners (FM and PI) while there is a similar but smaller trend with changing partners (RM). Without feedback (treatment NF) there are no significant changes in the level of reasoning reflected in belief statements.*

3.5 Consistency of actions and beliefs

We finally turn to the question whether actions and beliefs are consistent, i.e., whether a participant who holds a level-k belief also chooses a level-k action. For these consistency checks, we rely on the classifications of stated beliefs as noisy responses to the true underlying beliefs and assume that the parameter θ governing the belief and the action choice belongs to the same interval.[19]

In a first step we assess whether action and belief types are dependent. This can be tested for the two instances where action and belief types exactly correspond. First, for the row player the fraction of $L1$ actions given $L1$ beliefs is significantly higher than the fraction of $L1$ actions given

[18] The only qualitative differences to Table 6 are that the row players' weight trends of $L2^{+}$ in RM and $L4^{+}_{\theta\leq.15}$ in PI are insignificant. However, the non-parametric tests are based on crude classifications of stated beliefs as compared to the mixture model approach in (3) that accounts for the proximity of the stated beliefs to each underlying belief.

[19] If θ changes between the statement of the belief and the action choice within one period, we count this as an inconsistency. We also expect that subjects with a θ near the limits of the intervals for which the predicted action and belief change show more inconsistent behaviour if θ is not stable within periods.

Table 7: Consistency of actions and beliefs.

	Belief type (b)		Action(s) a consistent with b		Random $\Pr(a\|b)$	Observed $\Pr(a\|b)$	N
	(1)		(2)		(3)	(4)	(5)
Row player	$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$	$L1$	Middle	$L1$	0.333	0.348	824
	$(0,1,0)$	$L2^{+}$	Top/Bottom	$L2^{+}_{\theta\leq .15}/L2^{+}_{\theta>.15}$	0.667	0.681 †	774
	$(0,0,1)$	$L3^{+}_{\theta>.40}$	Bottom	$L2^{+}_{\theta>.15}$	0.333	0.901***§	204
	$(1,0,0)$	$L4^{+}_{\theta\leq .15}$	Top	$L2^{+}_{\theta\leq .15}$	0.333	0.873***‡	258
Column player	$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$	$L1$	Center	$L1^{+}$	0.333	0.607***§	850
	$(0,1,0)$	$L2$	Center/Right	$L1^{+}/L2^{+}_{\theta>.40}$	0.667	0.954***§	284
	$(0,0,1)$	$L3^{+}_{\theta>.15}$	Center/Right	$L1^{+}/L2^{+}_{\theta>.40}$	0.667	0.992***§	640
	$(1,0,0)$	$L3^{+}_{\theta\leq .15}$	Left	$L3^{+}_{\theta\leq .15}$	0.333	0.825***§	286

Notes: Column (1) states the four belief types for each player role and column (2) denotes the action(s) consistent with each belief type. Column (3) denotes the expected consistency when subjects' action choices are random. Column (4) shows the observed consistency. *** indicates significant differences to random choices (Wald tests, $p < 0.01$). §, ‡ and † indicate that comparisons are significant (two-sided Wald tests, $\alpha = 0.05$) in all treatments, in RM, FM and PI and, respectively, in FM only.

beliefs that are different from $L1$. Second, for the column player, the probability of observing the $L3^{+}_{\theta\leq .15}$ action given $L3^{+}_{\theta\leq .15}$ beliefs is significantly higher than observing the $L3^{+}_{\theta\leq .15}$ action given beliefs that are different from $L3^{+}_{\theta\leq .15}$.[20] Thus, actions and beliefs are significantly dependent.

In a second step we assess to what extent the subjects best respond to their beliefs, as predicted in standard Nash equilibrium and in the level-k model. We check whether the probability of choosing an action that is consistent with a given belief type is significantly different from random behavior. These conditional probabilities are provided in Table 7. For the row player, the fraction of $L1$ actions given $L1$ beliefs is merely 34.8%, which is not significantly different from random behavior. Also the row players' probability of choosing Top or Bottom given $L2^{+}$ beliefs around $(0,1,0)$ is not significantly higher than expected with random behavior. However, for all other possible comparisons both for the row and the column player, the conditional choice probabilities display a significantly higher consistency than random choices. For example, in 90.1% of the cases where row players hold $L3^{+}_{\theta>.40}$ beliefs close to $(0,0,1)$, they play Bottom (level $L2^{+}_{\theta>.15}$), which is significantly different from random behavior in each treatment and on the aggregate level.[21]

[20] We run a probit regression with an $L1$-action ($L3^{+}_{\theta\leq .15}$-action) indicator variable as the dependent variable and an $L1$-belief ($L3^{+}_{\theta\leq .15}$-belief) indicator variable as the independent variable in order to control for observational clusters on the individual level, which yields $p = 0.003$ ($p < 0.001$). Fisher's exact test also rejects the independence of both measures for both player roles at $p < 0.001$. On the treatment level, independence for the row player is rejected in FM and RM, while it is rejected in all treatments for the column player.

[21] We can compare our data to best-response rates found in similar studies by assuming $\theta = 0$ and calculating the best-response rate as the fraction of choosing the action with the highest expected payoff given one's belief as in the literature. This yields an overall best response rate of 63% for our data. In related experiments of normal-form games with constant-sum 2×2 games (Nyarko and Schotter, 2002) or with 3×3 constant-sum and variable-sum games

Finally, we observe inconclusive and largely insignificant results for the development of consistency over time.

Result 3 *Actions and beliefs as measures of strategic sophistication show a significant positive dependence. The probability of a subject choosing an action that is consistent with her belief type is significantly higher than with random behavior except for lower level beliefs (L1 or L2^{+}) of the row player.*

4 Conclusions

We performed an experiment to study the development of strategic reasoning over a finite number of periods. In order to understand the determinants of learning to play strategically, we varied the information available to the subjects. In addition to observing the action choices of the subjects, we elicited their beliefs about the opponent's actions. In the 3×3 normal-form game employed in our study, we used predictions of the level-k model pioneered by Nagel (1995) in order to classify actions and beliefs as choices by level-k types. For the type classifications, we allowed for heterogeneity of the subjects' preferences.

The results of our experiment show that subjects learn to play the game in the sense of reaching higher levels of reasoning in environments with feedback, independent of the information on the payoffs of the other player and the matching protocol. This validates the use of inductive learning models for the type of game we have studied. Nonetheless, our design of providing or withholding information about the game and the other players only allows us to study a subset of the complex phenomenon of human learning. Other forms of learning include, for example, social learning where people infer information from observing others' behavior (Bikhchandani et al., 1992) as well as learning through contemplation or reflection (Agranov et al., 2010), which can be seen as a form of deductive reasoning. Furthermore, learning in groups can differ from learning individually, e.g. teams have been found to be better at transferring knowledge across games than individuals (Cooper and Kagel, 2005). Psychologists have attempted to understand how human reasoning functions on a more fundamental level. For example, analogy-based thinking has been identified as the basis for induction (Holyoak, 1985) while the construction of mental models of a situation can explain how people make logical inferences and deductions (Johnson-Laird, 2009).

(Rey-Biel, 2009), best response rates are about 70%, whereas the rates range from 49% to 63% in the games used in Costa-Gomes and Weizsäcker (2008) and Hyndman et al. (forthcoming).

The main contribution of this paper is its use of the level-k model to track learning in actions and beliefs in the presence of heterogeneous preferences. This framework, which can also be applied to other games, allows us to show that actions and beliefs evolve in a parallel manner. Our finding of a co-evolution of actions and beliefs to higher levels of reasoning complements the findings of other studies about beliefs in repeated normal-form games. First, we have used stated beliefs in our analysis, based on the finding of Nyarko and Schotter (2002) that stated beliefs are better predictors of action choices than beliefs constructed from past observed actions, a result that we were able to replicate with our data.[22] Second, we observe that subjects choose an action that is consistent with their belief type for almost all types, but the consistency is lower for subjects with lower levels of reasoning. The results of Costa-Gomes and Weizsäcker (2008) provide a good benchmark regarding such inconsistencies of actions and beliefs in one-shot games, but it is by now very little understood under which conditions players best respond more to their beliefs than in others, whether this depends on their strategic sophistication, and under which conditions they can learn to be consistent. Although we took one step into this direction, in our view these questions deserve more thorough empirical scrutiny.

References

[1] Anderson, S.P., Goeree, J.K. and Holt C.A. (1998). A theoretical analysis of altruism and decision error in public goods games. *Journal of Public Economics*, 70, 297-323.

[2] Agranov, M, Caplin, A. and Tergiman, C. (2010). The Process of Choice in Guessing Games. Mimeo.

[3] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100, 992-1026

[4] Blanco, M., Engelmann, D., Koch, A. and Normann, H.-T. (2010). Belief Elicitation in Experiments: Is there a Hedging Problem?, *Experimental Economics*, 13, 412-438.

[5] Camerer, C., Ho, T., and Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games, *Quarterly Journal of Economics*, 119, 861-898.

[22] For the exact findings, see the working paper by Fehr et al. (2008). In contrast, Rutstrom and Wilcox (2009) find that inferred beliefs can be better than stated belief in a matching pennies game, especially when payoffs are highly asymmetric.

[6] Cooper, D. J. and Kagel, J. H. (2005). Are Two Heads Better than One? Team versus Individual Play in Signaling Games, *American Economic Review*, 95,477-509

[7] Costa-Gomes, M.A., Crawford, V. and Broseta, B. (2001). Cognition and Behavior in Normal-form Games: An Experimental Study, *Econometrica*, 69, 1193-1235.

[8] Costa-Gomes, M.A. and Weizsäcker, G. (2008). Stated Beliefs and Play in Normal-form Games, *Review of Economic Studies*, 75, 729-762.

[9] Cox, J. C., Friedman, D. and Gjerstad, S. (2007). A tractable model of reciprocity and fairness, *Games and Economic Behavior*, 59, 17 - 45.

[10] Edgeworth, F. (1881). *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: Kegan Paul.

[11] Hyndman, K., Özbay, E.Y., Schotter, A. and Ehrblatt, W. Z. (forthcoming). Convergence: An Experimental Study of Teaching and Learning in Repeated Games, *Journal of the European Economic Association*.

[12] Fehr, D., Kübler, D. and Danz, D. (2008). Information and Beliefs in a Repeated Normal-form Game, SFB 649 Working Paper No. 2008-026.

[13] Fischbacher, Urs (2007). z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10, 171-178.

[14] Goeree, J. K., Holt, C. A., Laury, S. K. (2002). Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior. *Journal of Public Economics*, 83, 255–276.

[15] Greiner, B. (2004). *An Online Recruitment System for Economic Experiments*, in: Kremer,K., Macho, V. (eds.): Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63, Göttingen: Ges. für Wiss. Datenverarbeitung, 79-93.

[16] Holyoak, K. J. (1985). *The Pragmatics of analogical transfer*. In G. H. Bower (Ed.), The psychology of learning and motivation, Vol. 19, 59-87. New York: Academic Press.

[17] Johnson-Laird, P.N. (2009). *Reasoning*, in Rabbitt, P. (Ed.) Inside Psychology: A Science over 50 Years. Oxford: Oxford University Press, 167-177.

[18] Kalai, E. and Lehrer, E. (1993). Subjective Equilibrium in Repeated Games, *Econometrica*, 61, 1231-1240.

[19] Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study, *American Economic Review*, 85, 1313-26.

[20] Nyarko, Y. and Schotter, A. (2002). An Experimental Study of Belief Learning Using Elicited Beliefs, *Econometrica*, 70, 971-1005.

[21] Oechssler, J. and Schipper, B. (2003). Can You Guess the Game You Are Playing?, *Games and Economic Behavior*, 43, 137-152.

[22] Rey-Biel, P. (2009). Equilibrium Play and Best Response to (Stated) Beliefs in Constant Sum Games, *Games and Economic Behavior*, 65, 572-585.

[23] Rick, S and Weber, R. (2010). Meaningful Learning and Transfer of Learning in Games Played Repeatedly Without Feedback, *Games and Economic Behavior*, 68, 716-730.

[24] Rutstrom, E. and Wilxoc, N. (2009). Stated Beliefs versus Inferred Beliefs: A Methodological Inquiry and Experimental Test. *Games and Economic Behavior*, 69, 616-632.

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

[26] Stahl, D.O. and Wilson, P.W. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*,10, 218-254.

[27] Terracol, A. and Vaksmann, J. (2009). Dumbing Down Rational Players: Learning and Teaching in an Experimental Game. *Journal of Economic Behavior and Organization*, 70, 54-71.

[28] Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, *Econometrica*, 57, 307-333.

[29] Weber, R. (2003). 'Learning' With No Feedback in a Competitive Guessing Game, *Games and Economic Behavior*, 44, 134-144.

Appendix

Sample Instructions (for FM)

The experiment you are about to participate in is part of a project financed by the German Research Foundation (DFG). Its aim is to analyze economic decision-making behavior. You can earn a considerable amount of money in this experiment, dependent on your decisions and the decisions of the other participants. Consequently, it is extremely important that you read these instructions very carefully.

Please note: these instructions are for your eyes only, and it is not permitted to hand on any information whatsoever to other participants. Similarly, you are not allowed to speak to the other participants throughout the whole experiment. Should you have a question, please raise your hand and we will come to you and answer your question individually. Please do not ask your question(s) aloud. If you break these rules, we will unfortunately be compelled to discontinue the experiment.

General information The experiment is made up of several periods where decisions must be made and questions answered. You can win points with your decisions. These points represent your earnings and will be converted into euros at the end of the game and paid out in cash. The exact procedure of the experiment, the various decisions and the method of payment are clearly explained in the next section.

The decision-making situation At the beginning of the experiment, you will be assigned by draw to another participant, randomly and anonymously. This allocation is maintained throughout the whole of the remaining experiment. The participant who has been assigned to you will be called "the other one" from now on.

In each period, you and the other one will be confronted with the same decision-making situation. Each time, you must choose between the three alternatives: "top", "middle", and "bottom".

Each of these three alternatives has been given three possible payoffs (as points). The other one must also decide between three alternatives ("left", "center" or "right"), and each of these alternatives has also three possible payoffs, as above. You will see the following input screen on the computer:

Your three alternatives, "top", "middle", and "bottom", are listed in the first column of the table. Next to your alternatives, you can see three boxes, each with two numbers. The subscript (lower) number is always your possible payoff. On the input screen illustrated above, the alternative "top" has been allocated the payoff of 78, 72 and 12, the alternative "middle" the payoff of 67, 59 and 78, and the alternative "bottom" the payoff of 21, 62 and 89. This means that should you decide on "top", for example, then your payoff is 78, 72 or 12 points. The payoff you actually receive depends on whether the other one selects "left", "center" or "right". Thus your payoff depends on your own decision as well as that of the other one. The superscript (raised) number in any box is always the possible payoff of the other one. For example, if the other one decides on "left", then his/her possible payoff points are 68, 52 and 11. This means, for example, that if you decide on "middle" and the other one decides on "right", your payoff is 78 points. The payoff for the other one is 49 points in this case.

The possible payoff points on the input screen above are therefore as follows:

You choose “top”; the other one chooses “left”:

Your payoff is: 78 points

The payoff for the other one is: 68 points

You choose “top”; the other one chooses “center”

Your payoff is: 72 points

The payoff for the other one is: 23 points

You choose “top”; the other one chooses “right”:

Your payoff is: 12 points

The payoff for the other one is: 20 points

You choose “middle”; the other one chooses “left”:

Your payoff is: 67 points

The payoff for the other one is: 52 points

You choose “middle”; the other one chooses “center”:

Your payoff is: 59 points

The payoff for the other one is: 63 points

You choose “middle”; the other one chooses “right”:

Your payoff is: 78 points

The payoff for the other one is: 49 points

You choose “bottom”; the other one chooses “left”:

Your payoff is: 21 points

The payoff for the other one is: 11 points

You choose “bottom”; the other one chooses “center”:

Your payoff is: 62 points

The payoff for the other one is: 89 points

You choose “bottom”; the other one chooses “right”:

Your payoff is: 89 points

The payoff for the other one is: 78 points

Please note that the possible payoff points for you and the other one remain the same in every period.

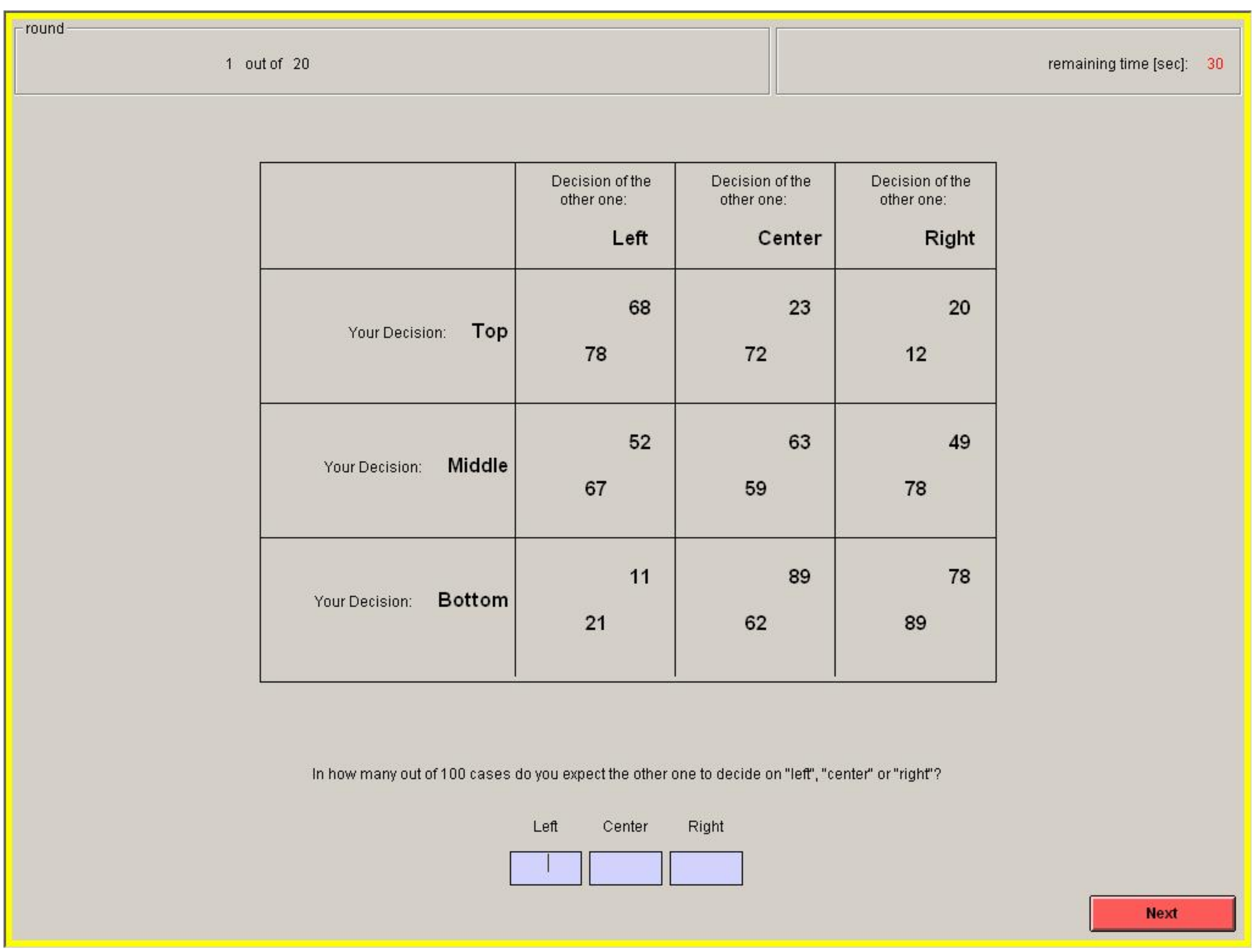
The other one always has exactly the same input screen in front of him/her as you do. After you and the other one have chosen between the three alternatives, you will be informed of your payoff in this period. This is the only information you will be given during the experiment in each period. The next period begins after that.

Statement of expectations

a) How can you state your expectations? Before each decision-making situation, you will be asked how you estimate the decision-making behavior of the other one. This means that at the beginning of each period we will require you to predict how the other one will decide in this period. You will have to answer the following question:

In how many out of 100 cases do you expect the other one to decide on "left", "center" or "right"?

Of course, the other one makes his decision only once in each period. You could also consider the question as asking you to state the likelihood that each of the three alternatives is chosen by the other one. You will see the following input screen on the computer:



Your three alternatives, “top”, “middle” and “bottom”, are listed in the table above, as well as the corresponding possible payoff. Below that, there is the question with the three boxes.

Let us assume that you are sure that the other one will choose “right”, and definitely not “center” or “left”. Then you would respond to our question by entering the number 100 in the box for “right” and the number 0 in the boxes for “center” and “left”. Alternatively, we could assume that you think the other one will probably choose “center”, but there is still a small chance that s/he will choose “right”, and an even smaller chance that s/he will choose “left”. Then, for example, you might respond to our question by entering the number 70 for “center”, 20 for “right” and 10 for “left”.

If you think it is even more unlikely that s/he will choose “center”, then you could enter, for example, 60 for “center”, 24 for “right” and 16 for “left”. Or it is possible that you think it is equally likely that the other one will choose “left”, “center” and “right”. Then you should enter, for example, the numbers 33, 33, 34 in the boxes.

Please note that the three numbers may not be decimal, and that they must always add up to 100.

N.B.: The numbers used in the examples have been chosen arbitrarily. They give you no indication how you and the other one decide.

b) How is the payoff for your stated expectations calculated? Your payoff is calculated after you have guessed how frequently the other one chooses his/her three alternatives. Your payoff depends on the difference between your estimate of the frequency of the decision and the actual decision made. Your payoff is higher when you have guessed that the other one often makes the “true” decision (which s/he really made), and it is lower when you have guessed that the other one will make this decision infrequently. Similarly, your payoff is higher when you have correctly predicted that the other one will not make a particular decision and then s/he in fact does not make the decision.

The exact calculation of the payoff is as follows: We calculate a number for each of the three alternatives. This number reflects how appropriate your estimate of the decision frequency of the corresponding alternative was. We take these three numbers to calculate your payoff.

First, we consider how well you predicted the alternatives which were actually chosen. Let us assume that the other one chose “left”. We then compare your estimate of how often the other one would choose “left” out of 100 cases with the number 100, and calculate the difference between the two. This difference is then multiplied by itself and the resulting number multiplied by the factor 0.0005. Thus, if you expected the other one to choose “left” in many out of 100 cases, then this number will be smaller (since the difference between your estimate and 100 is small) than if you expected that s/he would choose “left” in few out of 100 cases.

Then we consider how well you predicted that the other two alternatives would not be chosen. Let us assume again, for example, that the other one chose “left”, which at the same time means that “center” and “right” were not chosen. Then we take your estimate for the alternative “center” and multiply this by itself. The resulting number is again multiplied by the factor 0.0005. We apply this procedure again to your estimate for the alternative “right”. We then take the three numbers thus calculated and deduct them from the number 10. This determines the number of points you receive for your statement of expectations.

As an illustration of how your payoff might appear, let us consider three examples. Let us assume that the other one chose “left” and that your estimate for “left” was 100 and correspondingly 0 for the other two alternatives. This means that you have stated an estimate that is exactly right. Consequently, you earn the following points:

$$10 - 0.0005 * (100 - 100)^2 - 0.0005 * 0^2 - 0.0005 * 0^2 = 10$$

Let us assume again that the other one chose “left”. Your estimate for “left” was 60, for “center” 20 and for “right” 20, which means that your stated estimate predicted that the other one would choose “left” more frequently than “center” and “right”. Consequently, you earn the following points:

$$10 - 0.0005 * (100 - 60)^2 - 0.0005 * 20^2 - 0.0005 * 20^2 = 8.8$$

If we still assume that the other one chose “left”, but your estimate for “left” was 0, for “center” also 0 and for “right” 100, this means that your stated estimate was exactly wrong. Consequently, you earn the following points:

$$10 - 0.0005 * (100 - 0)^2 - 0.0005 * 0^2 - 0.0005 * 100^2 = 0$$

N.B.: The numbers used in the examples have been chosen arbitrarily. They give no indication how you and the other one decide.

These examples should make it clear that you will always receive a payoff of at least 0 points, and at most 10 points for your stated expectations. And the closer your estimations, the more money you earn. (You may be asking yourself why we have chosen such a payoff ruling as described above. The reason being that with such a payoff ruling, you can expect the highest payment when you state numbers that are closest to your own estimate.)

Procedure and payment The experiment consists of 20 periods altogether. In each period, you have to first state your estimate of the behavior of the other one, and then make your own decision.

At the end of the experiment, a period each for the decision-making situation and for the statement of expectations will be chosen randomly in order to determine your earnings in the experiment. The choice of both periods will be made randomly by the experiment leader throwing a dice. The chosen periods will then be entered onto the input screen by the experiment leader. At the end of the experiment, you will see an overview of your earnings from the decision-making situation and your earnings from the statement of expectation, as well as the total amount. The payoff that you have attained in the corresponding period chosen will be converted at a rate of

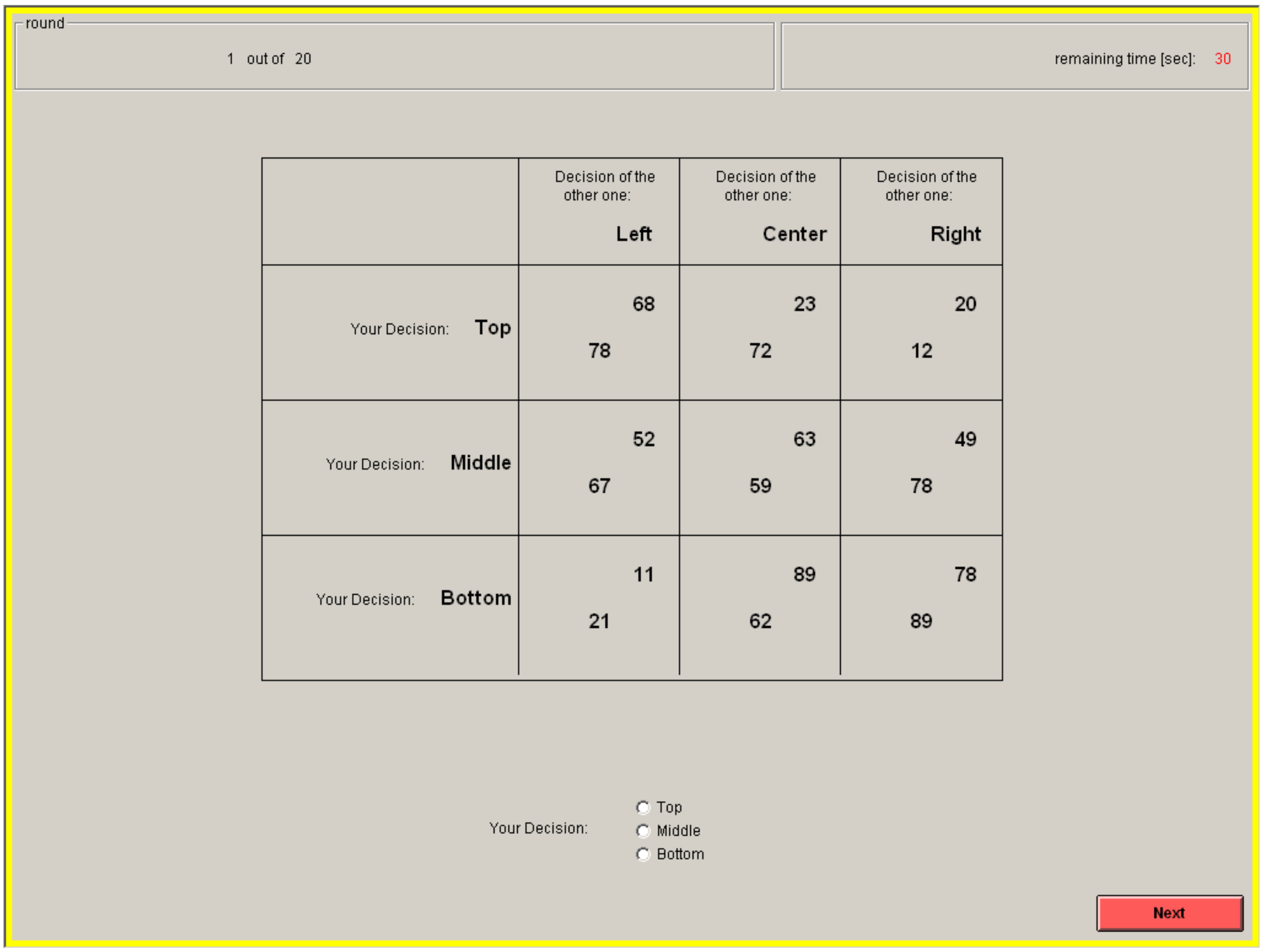
1 point = 15 cents

and will be paid out in cash.

Do you have any questions?

Control questions Now you have to answer 7 questions. In this way we are checking whether you have understood the decisions you have to make during the experiment. Should you have any further questions, please raise your hand and one of the experiment leaders will come to you. The experiment will not start until all participants have answered the control questions correctly.

The decision-making situation:



1. If you choose "bottom" and the other one chooses "center", how many points do you earn?

   _ _ _ _ _ _ _ _

2. If you choose "middle" and the other one chooses "left", how many points does the other one earn?

   _ _ _ _ _ _ _ _

3. If we assume your payoff amounts to 12, which decision did the other one make?

   _ _ _ _ _ _ _ _

4. If you choose “bottom” and the other one chooses “left”, how much do you earn and how much does the other one earn?

   The other one:_ _ _ _ _ _ _ _ _ _ _ _ You:_ _ _ _ _ _ _ _ _ _ _

5. Consider the following two cases:

   You expect the other one to choose “left” in 80 out of 100 cases. The other one actually does choose “left”. You expect the other one to choose “left” in 20 out of 100 cases. The other one actually chooses “right”. In both cases we assume that you expect the other one to choose “center” in 0 out of 100 cases.

   Is your payoff for the statement of expectation in the first case:

   higher the same lower (Please underline your answer!)

   than in the second case?

6. Imagine that Participant 1 states the following expectation: The other one chooses “left” in 50 out of 100 cases, “center” in 20 out of 100 cases, and “right” in 30 out of 100 cases. Participant 2 expects the following: the other one chooses “left” in 60 out of 100 cases, “center” in 20 out of 100 cases, and “right” in 20 out of 100 cases. We will assume that the other one chose “left” by Participant 1 as well as by Participant 2. Who will receive the highest payoff?

   Participant _ _ _ _ _

7. If you consider all three alternatives to be equally possible, which numbers should you then enter?

   left:_ _ _ _ _ _ _ _ _ center:_ _ _ _ _ _ _ _ _ _ right:_ _ _ _ _ _ _ _ _ _

Thank you for participating in the experiment!