Strategic sophistication of individuals and teams in experimental normal-form games*

Matthias Sutter#

University of Innsbruck and University of Gothenburg and IZA Bonn

Simon Czermak

University of Innsbruck

Francesco Feri

University of Innsbruck

Abstract

We present an experiment on strategic thinking and behavior of individuals and teams in one-shot normal-form games. Besides making choices, decision makers state their first-order beliefs about their opponent's choice, and second-order beliefs about the opponent's first order belief. We find that teams play Nash-strategies significantly more often, and their choices are more often a best reply to first order beliefs, something which teams also expect more often from their opponents. Using a maximum-likelihood error-rate model we find that teams can be classified as playing strategically in 62% of cases, while a majority of individuals (60%) plays non-strategically.

JEL-classification: C72, C91, C92

Keywords: Strategic thinking, beliefs, experiment, team decision making, individual decision making

This version: 14 November 2009

* We would like to thank Loukas Balafoutas, Martin Dufwenberg, Simon Gächter, Martin Kocher, Dorothea Kübler, John List and audiences at the University of Paderborn, the Spring Meeting of Young Economists in Lille and the ESA-European Meeting in Innsbruck for helpful comments. Financial support from the Max Planck Institute of Economics in Jena and the Austrian Science Foundation (through FWF-project P16617) is gratefully acknowledged.

Corresponding author. Address: Department of Public Finance, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. e-mail: matthias.sutter@uibk.ac.at

1. Introduction

Strategic sophistication refers to the extent to which players consider the structure of a game and the other players' incentives in the game before deciding on their strategy. There is a large literature that examines the strategic sophistication of individuals by means of experimental normal-form games (see, e.g., Stahl and Wilson, 1994, 1995; Haruvy, Stahl and Wilson, 1999; Costa-Gomes, Crawford and Broseta, 2001; Weizsäcker, 2003; Bhatt and Camerer, 2005; Costa-Gomes and Weizsäcker, 2008; Fehr, Kübler and Danz, 2008; Rey-Biel, 2009). In a nutshell, the main insight from this literature is the fact that strategic sophistication is often limited. For example, it has been shown that a considerable fraction of subjects ignores the incentives and the rationality of the other players, therefore showing non-strategic behavior (Weizsäcker, 2003). There is also a broad variety of strategic and non-strategic types, with boundedly rational strategic types being more frequent than types that play equilibrium consistently (Costa-Gomes et al., 2001). Interestingly, when eliciting beliefs of players about the other players' strategies, it turns out that in about 50% of cases players fail to best respond to their own stated beliefs (Costa-Gomes and Weizsäcker, 2008).[1] An explanation for the latter inconsistency has been identified by Costa-Gomes and Weizsäcker (2008) who have shown that subjects perceive a game differently when making choices than when stating beliefs (with stated beliefs revealing deeper strategic thinking than subjects' actions).

In this paper, we present an experiment in which we examine the strategic sophistication of teams (of three subjects each) and compare team decision making with individual decisions in 18 different, one-shot normal-form games that have been previously used to study individual strategic sophistication by Costa-Gomes et al. (2001). We are interested in the extent of strategic sophistication of teams in comparison to individuals. In particular, we compare the distribution of strategic and non-strategic types between teams and individuals and we examine whether teams are more frequently playing best response to their stated beliefs than individuals. We let teams and individuals make decisions in the 18 different games, and elicit their first order beliefs about their opponent's choice, and their second-order beliefs about the opponent's first order belief. Eliciting second-order beliefs – like in Bhatt

[1] Fehr et al. (2008) show that the frequency of best responses to own beliefs increases with repetition *if* feedback is given after each game. However, even under such conditions the relative frequency converges to only around 75% in the final round of their experiment (starting from slightly more than 50% in the first round, which is comparable to the level observed in experiments without feedback). Most of the other studies on strategic sophistication do not give feedback until the end of the experiment which prevents learning to best respond to own beliefs (as is also shown in a control treatment of Fehr et al., 2008).

and Camerer (2005), for example[2] – allows examining how players perceive the rationality of their opponents by checking how first- and second-order beliefs match. Hence, the analysis of beliefs is not confined to a player's own consistency of choice behavior and first-order beliefs. In our experiment we do not provide feedback until the very since we are mainly interested in the extent of strategic sophistication rather than learning. We let teams play against teams, and individuals against individuals, but also teams against individuals. The latter treatment allows examining whether individuals and teams discriminate in their behavior and beliefs with respect to the type of decision maker (individual or team) in the opponent's role. We classify teams and individuals by applying the maximum likelihood error-rate analysis of a decision maker's decisions that was developed by Costa-Gomes et al. (2001). The econometric model is a mixture model in which each decision maker's type is drawn from a common prior distribution over different strategic and non-strategic types and remains constant for all 18 games.

Generally speaking, investigating team decision making seems interesting and important because many economically important decisions are taken by teams rather than by individuals. One can think of company boards, management teams, committees, or central bank boards as relevant economic agents in a variety of contexts where strategic sophistication plays a role, such as in decisions on market entry, technology races, company takeovers, or how to optimally intervene through monetary policy instruments in financial markets during a crisis. In addition to this genuine relevance of team decision making, our paper contributes to the literature on strategic sophistication and the literature on team decision-making in the following ways.

We are not aware of any other study that examines the strategic sophistication of teams in normal-form games. While the basic bottom-line of the flourishing research on team decision making seems to be that teams are "more rational" than individuals in strategic games – meaning that team behavior is in the aggregate typically closer to standard game theoretic predictions than individual behavior (see, e.g., Bornstein and Yaniv, 1998; Bornstein, Kugler and Ziegelmeyer, 2004; Cooper and Kagel, 2005; Kocher and Sutter, 2005; Charness and Jackson, 2007)[3] – there is no paper on team decision making that classifies single teams as of a particular strategic or non-strategic type and compares the distribution of types across

[2] Note that Costa-Gomes and Weizsäcker (2008) mention to have also run some treatments where they had asked for second-order beliefs, but they do not report these treatments and their results in the paper.

[3] The paper by Cason and Mui (1997) is often misinterpreted as showing that teams are more generous than individuals in a dictator game. However, Cason and Mui (1997) did not find that teams in general are more generous than individuals, but only reported more other-regarding team choices when team members differed in their individual dictator game choices.

individuals and teams. The closest paper with respect to classifying teams as more or less strategic is by Cooper and Kagel (2005) who investigate individual and team decision-making in a signaling game where a market incumbent has to decide on a production quantity (as a signal for its type) before a potential market entrant decides on whether or not to enter the market. Cooper and Kagel (2005) find that teams are more frequently playing strategically by choosing higher production levels to signal a low-cost type market incumbent, thus making it less likely for outsiders to enter the market. While their study – in particular the excellent analysis of chat communication in the two-person teams – adds valuable insights about the strategic thinking of teams, it does not provide a detailed classification of types on the level of individual teams as will be done in our paper. Furthermore, we will analyze strategic sophistication of teams by considering not only choices, but also their first- and second-order beliefs. Beliefs have not been elicited in the study of Cooper and Kagel (2005). Eliciting beliefs will allow us to check whether teams are more likely to best reply to their own first-order beliefs, and whether teams expect their opponents to be best responding as well (by matching first-order beliefs with second-order beliefs). None of this has been studied systematically and in depth in the team decision-making literature before. Our paper can therefore provide a fine-grained picture of the (bounded) rationality and strategic sophistication of teams and how this compares to individual decision making.[4]

Based on a total of 192 experimental participants, we find that teams play Nash-strategies significantly more often, and their choices are more often a best reply to first order beliefs. Teams also expect their opponents to be more consistent, i.e. first-order beliefs are more often a best reply to second-order beliefs. Interestingly, neither teams nor individuals discriminate in their behavior and beliefs with respect to the type of decision-maker in the opponent's role. This implies that teams expect (mistakenly) their opponents to be more consistent in general, and not only in situations where a team faces another team in the opponent's role. Using a maximum-likelihood error-rate model we find that teams can be classified as strategic in about 62% of cases. The model type of team decision making is a *D1*-type which applies one step of deleting strategies that are dominated by pure strategies and then plays best reply to a uniform prior over the opponent's remaining strategies. Individuals are classified as strategic in only 40% of cases. Their modal type is a non-strategic, *naïve* type (which best responds to beliefs that assign equal probabilities to the opponent's available strategies).

[4] We will not address the issue of learning in team decision making, since this has been studied previously. The interested reader is referred to Kocher and Sutter (2005) or Feri, Irlenbusch and Sutter (2009) on how teams learn (in beauty-contest games and coordination games, respectively).

The rest of the paper is organized as follows. In section 2 we present the experimental design, including the normal-form games used in the experiment, the experimental treatments and procedure. Section 3 reports the experimental results. It starts with an aggregate analysis of decisions, first- and second-order beliefs of individuals and teams. Then we examine the consistency of players' choices with own first-order beliefs, and the players' perceived consistency of their opponents by matching first- and second-order beliefs. We continue with a regression that determines the factors that lead players to play the Nash strategy and to expect their opponent to play Nash as well. In section 4 we use a maximum-likelihood error-rate model to classify individuals and teams into either of eight different strategic and non-strategic types. Section 5 concludes the paper.

2. Experimental design

2.1. The 18 normal-form games

Figure 1 presents the 18 normal-form games that we used in our experiment. They are taken from Costa-Gomes et al. (2001), because these games are very well designed to study strategic thinking and strategic behavior. Out of the 18 games, 16 games are pairs of isomorphic games (which are identical for row and column players except for transformation of player roles and small, uniform payoff shifts). We applied the same order of games as in Costa-Gomes et al. (2001). This is indicated in Figure 1 as "game # x", with $x \in \{1, \ldots, 18\}$. In brackets after the game number we refer to two different types of games. "***D***"-games are those in which one of the two players has a strictly dominant strategy. In total, there are 10 ***D***-games among the 18 normal-form games. Note that in each of these 10 games the Nash equilibrium is Pareto-dominated by another strategy combination of row and column players. In "***ND***"-games no player has a strictly dominant strategy. Like in the ***D***-games, however, the Nash-equilibrium in all 8 ***ND***-games is Pareto-dominated by another strategy combination.

Figure 1 about here

The number of rows (columns) indicates the number of available strategies for row (column) players. The 18 games have different levels of complexity, interpreted as the required number of rounds of iterated pure-strategy dominance a [Row, Column] player needs to identify the own equilibrium choice, where [∞R, ∞R] denotes non-dominance solvable games. In ***D***-games, the number of rounds needed to reach the equilibrium choice is either 1 or 2. In ***ND***-games, players may need 2, 3, or an infinite number of rounds of iterated pure-strategy dominance to identify the equilibrium choice.

2.2. Treatments, decisions and payment

We have implemented three different treatments which are distinguished by which types of decision makers are interacting with each other in the normal-form games.

- **Ind**: In the individual treatment both row and column players were individuals.
- **Team**: In the team treatment both row and column players were teams of three subjects each. Team members sat together in a sound-proof cabin and could communicate with each other in order to reach a team-decision. There were no restrictions on communication and no regulations how a team should reach an agreement, except that only one decision could be entered by a team on the computer.
- **Mixed**: In the mixed treatment a team of three subjects was interacting with an individual. In half of the **Mixed**-sessions the team was in the role of the row-player and the individual in the role of the column-player, while the reverse applied in the other half of sessions.

Note that the treatment – and thus the type of decision maker in the opponent's role – was common knowledge in the experiment (see the experimental instructions in the Supplementary Material). To avoid any influence of the kind of presentation on the data, we presented all games to all participants in a way that they saw themselves as a row player. In each single game each player (either individual or team) had to make three different decisions:

- **Choice**: This was a player's choice of a row in a given game.
- **First-order belief (FOB)**: Each player had to indicate the action expected from the opponent. From the viewpoint of the player this meant to select a column as the expected choice of the opponent.
- **Second-order belief (SOB)**: Players stated a belief about their opponent's first-order belief. This meant to indicate a row that matches the player's belief about the opponent's expectation about the player's choice.

Each player had to make all three decisions in each of the 18 games, yielding a total of 54 decisions to be taken during the experiment. We proceeded game by game with all three decisions, but the order of the three decisions for a given game was randomly determined.[5] In order to suppress learning, partners changed after each game (random matching) and there was no feedback until the experiment was finished.

Payments in the experiment were determined as follows: At the end of the experiment each player needed to draw a card from a deck of cards which were showing numbers from 1 to 18. This card determined the game that was relevant for payment. Players then received full information about their own and their opponent's decision in this game and about the beliefs that they as well as their opponent had expressed. After that players were asked to draw another card from a deck of cards showing numbers 1 (for choice), 2 (for first-order belief), or 3 (for second-order belief). If a player was paid for her choice, the player received 30 Euro Cents for each experimental point earned in the game as a consequence from the own and the opponent's choice (as shown in Figure 1). Players who were paid for their first-order or second-order belief received 15 Euros if their belief was correct, but nothing if it was wrong. Note that in order to keep the per-capita incentives constant for individuals and teams the payoffs as described here were paid to each member of a team, and participants knew this.

The all-or-nothing feature for paying beliefs was intended to make the decisions on first- and second-order beliefs as salient as possible. Our approach with respect to eliciting and paying beliefs was also chosen for practical reasons, though. Given the relatively high number of decisions (54) to be taken in the experiment we opted for eliciting point beliefs rather than asking for a probability distribution of beliefs over all available strategies. In particular asking teams for a probability distribution would have been very time-consuming and would probably have exhausted participants' attention and patience in an experiment that already took on average 2 hours with our procedure.[6]

[5] In the experiment of Costa-Gomes and Weizsäcker (2008) the order of making choices and stating first-order beliefs was varied. Since they find only minor order effects we decided to order choices, first-order beliefs and second-order beliefs randomly for each game. However, all three decisions in a particular game were taken before proceeding to the next game.

[6] Haruvy (2002) or Costa-Gomes and Weizsäcker (2008) elicited (more informative) probability distributions of the opponent player's choices (i.e., a player's first-order beliefs). While in the study of Haruvy (2002) the computer then calculated the best response for the player's first order belief and implemented the best response automatically as the choice of the player, the subjects in the study of Costa-Gomes and Weizsäcker (2008) could make their choices and state their first-order beliefs independently. Our approach of asking for a single strategy as a belief and rewarding beliefs if they are exactly right induces subjects to report the mode of their (implicit) distribution of probabilities for the different available strategies.

2.3. Experimental procedure

In each treatment we conducted four sessions that were run in the video-lab of the Max Planck Institute of Economics in Jena.[7] This lab has eight sound-proof cabins which are connected via a computer network (using z-Tree by Fischbacher, 2007). Hence, we could invite eight decision makers for each session. Four decision makers were row players and four decision makers were column players. Depending upon the treatment, a decision maker was either an individual subject or a team of three subjects. When recruiting subjects with the help of ORSEE (Greiner, 2004), assignment to treatments was randomly determined. In the treatments with team decision making, subjects were randomly assigned to a particular team at the beginning of the experiment. Teams stayed together for the entire experiment. The matching of row and column players was random and changed after each game. Note that for a particular game (with its three decisions) the matching was kept constant. To make sure that the participants understood the instructions and the experimental tasks, they needed to answer a questionnaire before the experiment in which they were tested how choices and first- and second-order beliefs mapped into outcomes and payoffs.

In sum, we had a total of 192 participants, recruited from undergraduate students of the University of Jena. For each game in each treatment we obtained 32 different observations. No subject was allowed to participate in more than one session, and no participant had any prior knowledge in game theory. Sessions lasted on average 2 hours, and subjects earned on average 16.5 Euro, including a show-up fee of 4 Euro.

3. Experimental results

3.1. Pooling of data

In the following analysis we pool the data of row and column players, because 16 out of 18 games are isomorphic games which are practically identical for row and column players expect for small payoff shifts and the transformation of player roles (see Costa-Gomes et al., 2001, for details). All analyses presented below are based on all 18 games.

We have also examined whether players' choices and beliefs depend on the type of decision maker they face as their opponent. In other words, we tested whether the distribution

[7] Participants were not videotaped and they were told so. Individual decision makers were also seated alone in soundproof cabins in order to keep the conditions of individual and team decision making as identical as possible. We decided not to videotape team discussions because we wanted to minimize the possibility that differences between individuals and teams might be driven by the fact that teams were taped and thus observed.

of choices and first- and second-order beliefs differs for individuals between treatments **Ind** and **Mixed**, respectively for teams between treatments **Team** and **Mixed**. In none of these distribution tests we find any significant difference. This result indicates that decision makers do not condition their behavior and their beliefs on the type of decision maker in the opponent role. We summarize this as our first result:

Result 1: Neither individuals nor teams condition their choices and beliefs on the type of decision maker (individual or team) in the opponent role.

The first result also allows us to pool individual decisions from the treatments **Ind** and **Mixed**, as well as the team decisions taken in either **Team** or **Mixed**. Hence, when we talk of individual (team) decisions in the following, we refer to all decisions made by individuals (teams) in the experiment. Whether individuals and teams make different choices and state different beliefs (independent of the type of opponent they are facing) is subject of the following subsections.

3.2. Choices, first- and second-order beliefs of individuals and teams

In Table 1 we report the relative frequency of different strategies, taken either by teams or individuals across all games, and separately for the ***D***-games and the ***ND***-games.[8] We consider the strategy that leads to the Nash-equilibrium ("Nash"), the strategy that would yield an outcome that Pareto-dominates the Nash-equilibrium ("Pareto"), and other strategies ("Other"). The upper panel reports relative frequencies for own choices, the middle and lower panel for first-order, respectively second-order, beliefs.

Table 1 about here

We compare the relative frequencies of the different strategies chosen by individuals and teams by considering for each single decision maker the relative frequency of choosing a particular strategy across all games (or across ***D***- and ***ND***-games) and then apply a two-sided Mann-Whitney U-test to the resulting data. We see from the upper panel of Table 1 that Teams are significantly more often playing the Nash-strategy than individuals (49.30% vs. 40.97%; $p < 0.05$). Individuals choose more often what we call the "Pareto"-strategy of trying

[8] Tables S1 to S6 in the supplementary material accompanying this paper present the relative frequencies of each single strategy for each single game, both for individuals and teams, as well as for choices, first- and second-order beliefs.

to improve on the Nash-equilibrium choice (53.47% vs. 45.72%; $p < 0.05$). However, these differences are driven by the 10 games with a dominant strategy. Only in ***D***-games teams are significantly more often playing Nash, and less often playing Pareto, while there is no significant difference between individuals and teams in the games where neither player has a dominant strategy. This is a first indication that the complexity of a game has an impact on how the type of decision maker influences behavior. Table 2 will shed more light on this issue.

The middle panel in Table 1 considers first-order beliefs. Although teams expect their opponents to play Nash more often, respectively Pareto less often, than individuals, none of these differences is significant according to conventional levels. Only when considering second-order beliefs (in the lower panel) we find for the set of ***D***-games again the same type of differences between individuals and teams than we do for choice-data.[9] Actually, teams think that their opponents expect them to play Nash more often, respectively Pareto less often, than individual decision makers think about their opponents' expectations.

Comparing across panels in Table 1, while holding the type of decision maker constant, reveals that the relative frequency of playing Nash decreases from top to bottom, i.e. is highest for own choices, intermediate for first-order beliefs and lowest for second-order beliefs. This holds true both for individuals and teams and is significant for both types of decision makers ($p < 0.05$; Page tests). The reverse pattern applies for (potentially) Pareto improving strategies. Hence, equilibrium play is less likely according to beliefs than according to own choices, and less likely the higher the order of beliefs.

Table 2 about here

In Table 2 we examine the influence of a game's complexity on choices and beliefs in more detail. We present the relative frequencies of chosen strategies separately for games with and without a dominant strategy (***D***- vs. ***ND***-games) and for each type of game we distinguish according to the different number of rounds of iterated pure-strategy dominance a player needs to identify the own equilibrium choice. The first column shows that in games with a dominant strategy those players with the dominant strategy play Nash in about 80% of

[9] In the supplementary material we provide in Tables S7 to S10 the relative frequency with which individuals and teams were correct in their first- and second-order beliefs. On average, teams were marginally more often correct in their beliefs (62.7% vs. 61.0% for first-order beliefs; 69.1% vs. 68.6% for second-order beliefs), but the difference is far from being significant.

cases. There is no significant difference between individuals and teams, neither for choices nor for first- or second-order beliefs.

However, the second column reveals a very strong difference between individuals and teams when they are in the role of the player without the dominant strategy (when the opponent has one). While teams choose the Nash-strategy in almost 60% of cases then, individuals play Nash in less than 40%. The reverse holds true for the Pareto strategy. Given the high probability of players with a dominant strategy to actually choose the dominant strategy (see first column), the behavior of teams in column 2R(D) can be judged as more rational, since playing what we call the Pareto strategy when the opponent plays Nash would lead to lower payoffs than when playing Nash. These differences in choices between individuals and teams when they need two rounds of iterated pure-strategy dominance to identify the own equilibrium choice is again mirrored in second-order beliefs, albeit at a much lower level. Teams also think that their opponents expect them to play Nash almost double as often as individuals think of their opponents (25.42% vs. 13.75%). The differences in first-order beliefs are qualitatively similar, but fail significance at conventional levels.

The third column in Table 2 considers games without a dominant strategy for either of the players, and it clearly confirms the findings from the second column. When two rounds of iterated dominance are needed, teams play significantly more often Nash – and less often Pareto – than individuals. The same pattern is found for first- and second-order beliefs (with the differences being weakly significant for first-order beliefs). The final two columns in Table 2 show that for more complex situations (with more than two rounds of iterated dominance) there are no significant differences in strategy choices between individuals and teams. Thus, the evidence from Table 2 suggests that for very straightforward decisions (when a dominant strategy is given) and for more complex decisions (with more than 2 rounds of iterated dominance) there are no differences between individuals and teams. Yet, for decisions with 2 rounds of iterated dominance teams choose (and think that they are expected to choose) more often the equilibrium-strategy. We summarize the findings in this subsection in our next result:

Result 2: Overall, teams play the equilibrium strategy of the 18 games significantly more often than individuals. Taking a game's complexity into account we find that individuals and teams do not differ in the likelihood of playing equilibrium when they have a dominant strategy or need more than two steps of iterated dominance. In situations where two steps are required, though, teams play equilibrium much more often. These differences in choice-

patterns of individuals and teams are also reflected in their beliefs, albeit not always significantly.

3.3. Consistency of choices and beliefs

We classify a player's choice as consistent with beliefs, if the choice is a best reply to the player's own first-order belief. Likewise, we define the opponent's expected consistency as to whether the first-order belief is a best reply to a player's second-order belief. The latter determines basically a player's first-order belief about the opponent's consistency.

Table 3 about here

Table 3 reports in the upper part the players' own consistency and in the lower part the expected consistency of the opponent player. Consistent with earlier findings by Costa-Gomes and Weizsäcker (2008) or Fehr et al. (2008) we find that only about one half of the individual decisions (55.79%) are a best-reply to the individual's own first-order (point) belief. However, we find that teams play best-reply significantly more often (65.86%). Again we see from the column ***D***-games that this difference is predominantly driven by the games where one party has a dominant strategy (and where the main difference is again due to the player who does not have the dominant strategy – more detailed results are available upon request). The lower part of Table 3 indicates that teams also expect their opponents on average to be more often consistent by playing best reply. However, these differences are not significant at conventional levels. We summarize these results in

Result 3: Teams make more often consistent choices than individuals, such that the actions of a team are more frequently a best reply to their first-order beliefs about the opponent's behavior. The level of consistency observed for individuals confirms earlier experimental findings. Hence, team decision making is a means for getting more consistent behavior.

3.4. A closer look at the consistency of choices and beliefs

In this subsection we want to examine the consistency of choices and beliefs in more detail by having a look at the two most frequent types of consistency. One type is a straightforward combination of playing Nash and expecting the opponent to play Nash as well. We call this the "Nash-consistency" (Nash-CON). The other very frequent type is when the combination of a player's choice and first-order belief would yield the maximal payoff that is available for the player in a specific game. We call this the "Maximum-consistency" (Max-CON). Note that Max-CON cannot coincide with the Nash-equilibrium in our games. Of course, the relative frequency of these two consistency types can be examined for a player himself, but also for the player's expectation about the opponent's consistency type.

Table 4 about here

Table 4 reports the relative frequency of Nash-CON and Max-CON. For the player's own consistency (see upper part) we find a very strong difference between individuals and teams with respect to Nash-CON. Again, the difference is mainly driven by our ***D***-games. While only 27% of individual choices and beliefs can be classified as Nash-CON, it is 45% of team choices and beliefs in the 10 games where one player has a dominant strategy ($p < 0.05$). Interestingly, for individuals it is even the case that Max-CON is (marginally) more frequent than Nash-CON, while for teams the ratio is almost 1:2. From the lower part of Table 4 we note that teams expect their opponents significantly more often to be of the Nash-CON-type than individuals, while individuals expect opponents more often to be of the Max-CON-type.

Result 4: Teams play the Nash-strategy with the expectation that the opponent also plays Nash much more often than individuals do. Individuals' choices and beliefs more often focus (optimistically) on the maximum available payoff in a given game. Teams also expect their opponents more often to play and expect Nash (first- and second-order beliefs) than individuals do.

3.5. The determinants of consistency-type Nash-CON

As a final step in analyzing the consistency of decisions we examine the determinants that increase the likelihood of the "Nash-consistency". Note that this is precisely the type of consistency standard game theory would predict for rational and payoff-maximizing decision makers. Hence, estimating the factors that make "Nash-consistency" more likely does not only reveal which factors promote this type of consistency, but also which ones prevent decision makers from playing the equilibrium strategy and expecting their opponent to do the same. In Table 5 we report a probit estimation of Nash-CON on various factors that are explained in the following. Since each decision maker had to make decisions in 18 different games, we cluster the standard errors on the decision maker.

Table 5 about here

Panel [A] of Table 5 shows that the consistency Nash-CON is more likely if players have a dominant strategy themselves ("dominant strategy") or if their opponent has one ("opponent dominant strategy"). The next three independent variables are all insignificant. Making decisions in a team ("team player"), having a team as opponent player ("opponent = team") or playing as a team against another team ("team * opponent = team") does not influence the likelihood of "Nash-consistency". The latter two terms confirm that the type of decision maker in the opponent role does not matter for decision making. However, the interaction of team decision making and having a dominant strategy ("team * dominant strategy") or the opponent having a dominant strategy ("team * opponent dominant strategy") has a significantly positive effect each. This is in line with our previous findings that decisions in the ***D***-games mainly drive the differences between individuals and teams. The latter two independent variables show that team decision making in games where a dominant strategy is available increases the likelihood of decisions and beliefs to be in accordance with standard game theoretic decisions.

The final three variables in panel [A] of Table 5 consider the magnitude of payoffs in a game in different ways. The variable "risk with maxstrategy" measures the absolute difference between the decision maker's payoff in the Nash-equilibrium and the decision maker's payoff in case he chose the strategy in which the game's highest payoff is possible while the opponent played Nash. In other words, this variable indicates how much money is at risk when a decision maker tries to reach his maximal possible payoff instead of playing Nash. The estimation shows that Nash-CON gets weakly significantly more likely the more money is at risk when deviating to the strategy that could maximize one's payoff. The next variable ("inequality in Nash") measures the absolute difference of both players' payoffs in the Nash equilibrium. The higher this difference, the less likely is consistent behavior of the Nash-CON-type. This means that the Nash-equilibrium and the expectation of the other player choosing Nash also gets less attractive when the Nash-equilibrium is associated with increasing inequality of payoffs. Hence, distributional preferences influence the consistency of choices and first-order beliefs. The final variable "diff. to opp. payoff in Nash" measures the difference between the decision maker's payoff and the opponent's payoff in the Nash equilibrium. A positive (negative) value indicates that the player earns more (less) than the opponent. Higher values make Nash-consistency more likely. This means also that if a player's payoff becomes relatively smaller in comparison to the opponent's payoff then the player is less likely to play Nash and expect the opponent to play Nash.[10]

Panel [B] of Table 5 reports the marginal effects of team decision making (compared to individual decision making) for specific combinations of the binary variables from panel [A], evaluated at the means of the three ultimate variables in panel [A]. If a decision maker has no dominant strategy, while the opponent has one, the likelihood of observing Nash-consistency is more than 20% higher when teams instead of individuals make choices and state their first-order beliefs. If a decision maker has a dominant strategy, while the opponent doesn't, then teams are also more likely to be Nash-consistent, however the effect fails significance.

Table 6 about here

Table 6 reports a similar regression as Table 5, but takes as the dependent variable a player's expectation about the opponent being a Nash-CON-type. With the exception of two variables ("risk with maxstrategy" and "difference to opponent's payoff in Nash") we find the same significant relationships as in Table 5, for which reason we do not repeat the description of the directional influence of the independent variables on the likelihood to expect the opponent to be of the Nash-CON-type here once more. We summarize the findings in

Result 5: Teams are more likely than individuals to be a Nash-consistency type – and expect their opponents to be of this type also more often – when the game has a dominant strategy. Inequality in payoffs in the Nash equilibrium makes consistent decisions (of choices and first-order beliefs, respectively of first- and second-order beliefs) less likely, both for individuals and teams. Hence, distributional preferences affect the degree of standard game-theoretic rationality.

4. Estimation of strategic and non-strategic types

In this section we present a maximum likelihood error-rate analysis of agents' decisions following the econometric model used in Costa-Gomes et al (2001). This econometric model is a mixture model in which each agent's type is drawn from a common prior distribution over

[10] Interacting any of the final three variables with the variable "team player" does not show any significant effects.

eight types and remains constant for all 18 games. The eight types that we consider can be classified into non-strategic and strategic types and they are defined as follows[11]:

Non-strategic types: (1) An *altruistic* type tries to maximize the sum of payoffs to himself and the opponent, implicitly assuming that the opponent is also *altruistic*. (2) A *pessimistic* type chooses the strategy that secures the best of all worst outcomes. Hence, a *pessimistic* type plays maximin. (3) A *naïve* type assigns equal probabilities to the opponent's strategies and best responds to this naïve belief. Note that a *naïve* type is equivalent to an *optimistic* one in our games, since an optimistic type plays maximax by maximizing the maximum payoff over the opponent's decisions.

Strategic types: (4) Type *L2* plays best response to *naïve* types. (5) Type *D1* applies one step of deleting strategies that are dominated by pure strategies and then plays best reply to a uniform prior over the opponent's remaining strategies. (6) Type *D2* applies two steps of deleting dominated strategies and then best responds to the opponent's remaining strategies. (7) An *equilibrium* type makes equilibrium choices (which are unique in our games). (8) A *sophisticated* type plays best reply to the probability distribution of opponents' strategies by taking the actually observed distribution of strategies in the experiment's subject pool as the estimated probability distribution.[12]

For the estimation of the mixture model let $i = 1, \ldots, N$ index the different agents (individuals or teams), let $k = 1, \ldots, K$ index our types, and let $c = 2, 3$, or 4 be the number of an agent's possible decisions in a given game. We assume that a type-k agent normally makes type k's decision, but in each game he makes an error with probability $\varepsilon_k \in [0,1]$, type k's *error rate*, in which case he makes each of his c decisions with probability $1/c$. For a type-k agent, the probability of type k's decision is then $1-\frac{(c-1)}{c}\varepsilon_k$. So the probability of any single non-type k decision is $\frac{\varepsilon_k}{c}$. We assume errors are independently and identically distributed across games and agents.

The likelihood function can be constructed as follows. Let T^c denote the total number of games in which agents have c decisions; in our design we have T^2 =11, T^3 =6, and T^4 =1.

[11] We follow Costa-Gomes et al. (2001) in the selection of types to be considered. As they indicate, the definition of types is largely based on earlier work by Stahl and Wilson (1994, 1995). Note that Costa-Gomes et al. (2001) introduce nine types. However, naïve and optimistic types cannot be distinguished in the games used in Costa-Gomes et al. (2001) and here.

[12] The best responses for each type are indicated in Figure 2 of Costa-Gomes et al. (2001, p. 1203). Of course, the best response of *sophisticated* types depends on the actually observed probability distribution of strategies in our experiment, and need not coincide with the strategy indicated in Figure 2 of Costa-Gomes et al. (2001).

Then let x_k^{ic} denote the number of agent *i*'s decisions that equal type *k*'s in games in which he has *c* decisions with $x_k^i = (x_k^{i2}, x_k^{i3}, x_k^{i4})$, $x^i = (x_1^i, ..., x_K^i)$, and $x = (x^i, ..., x^N)$. Let p_k denote agents' common prior *k*-type probability, and $\sum_{k=1}^{K} p_k = 1$ and $p = (p_1, ..., p_K)$. Let ε_k denote the *k*-type error rate and $\varepsilon = (\varepsilon_1, ..., \varepsilon_K)$. Given that a game has one type-*k* decision and *c*-1 non-type-*k* decisions, the probability of observing a particular sample with x_k^i type-*k* decisions when agent *i* is type *k* can be written as:

$$L_k^i(\varepsilon_k | x_k^i) = \prod_{c=2,3,4} \left[1 - \frac{(1-c)}{c}\varepsilon_k\right]^{x_k^{ic}} \left[\frac{\varepsilon_k}{c}\right]^{T^c - x_k^{ic}} \tag{1}$$

Weighting the right-hand side by p_k, summing over *k*, taking logarithms, and summing over *i* yields the log-likelihood function for the entire sample:

$$\ln L(p, \varepsilon | x) = \sum_{i=1}^{n} \ln \sum_{k=1}^{K} p_k L_k^i(\varepsilon_k | x_k^i) \tag{2}$$

We provide two separate estimations of (2), one with individuals as decision making agents and one with teams as agents. Table 7 presents the main estimations results for individual choices (left-hand side) and team choices (right-hand side).

Table 7 about here

In the analysis of the results we interpret the estimated probability p_k as the probability to find an agent of type *k* in the population under observation. From Table 7 we can see that three types have a significantly positive probability, both for individuals and teams. These are the *altruistic* type, the *naïve* type, and type *D1*. While the probability to find an *altruistic* type does not differ between individuals and teams, there are clear differences concerning the other two significant types. Individuals are significantly more often of the *naïve* (i.e. non-strategic) type than teams. On the contrary, teams are significantly more likely to be of the *D1*-type, which is a strategic type. Taking together all strategic types, we find that individuals are estimated as strategic types in only 40% of cases, while 62% of teams are classified under any of the strategic types.[13]

Result 6: The likelihood of teams to be any of five different strategic types (equilibrium, sophisticated, D1, D2, L2) is about 62%, while it is only 40% for individuals. The latter are predominantly acting non-strategically (as altruistic, pessimistic, or in particular naïve types).

5. Conclusion

In this paper we have studied strategic sophistication of individuals and teams in 18 different normal-form games. We have found that teams are more likely to play these games strategically – meaning that they consider in their choices the structure of the game and the incentives of their opponents when making decisions – than individuals do. According to the econometric estimation presented in section 4 we have been able to estimate the likelihood of strategic play for teams to be 62%, while it is only 40% for individuals. The modal type of team decision making is a *D1*-type which applies one step of deleting strategies that are dominated by pure strategies and then plays best reply to a uniform prior over the opponent's remaining strategies. On the contrary, the modal type of individual decision making is the *naïve* type which assigns equal probabilities to the opponent's strategies and best responds to this naïve belief. The *naïve* type coincides in our games with what is called the *optimistic* type in Costa-Gomes et al. (2001). An *optimistic* type expects to get the maximal payoff possible in a game.

A detailed analysis of individual and team behavior across games in section 3 has shown that teams play the Nash equilibrium significantly more often than individuals. Interestingly, this difference between individuals and teams is mainly driven by games with an intermediate degree of complexity. When games are very straightforward (when the decision maker has a dominant strategy) or when they are relatively complex (by requiring more than 2 steps of iterated dominance to figure out the equilibrium choice) there is no difference between individuals and teams.

Matching choices and first-order beliefs we have found that teams are more likely to play a best reply with their choices on their own first-order beliefs. While this kind of consistency

[13] At first sight it might seem surprising why the *equilibrium* type is so rare, while Table 1 has shown about 50% of equilibrium play by teams across all games. For the classification of types it is important what a decision maker chooses when, for example, *D1* and *equilibrium* do not suggest the same choice (as is the case in 6 of our games). In these games, the large majority of teams opts for the *D1*-strategy rather than the *equilibrium*-choice.

of choices and beliefs is only found in 56% of individual decisions, it prevails in 66% of team decisions. It seems noteworthy that the relative frequency of consistency of individual decisions is very similar to the one reported in a recent study by Costa-Gomes and Weizsäcker (2008). Our findings, however, indicate that this relatively low level of consistency can be significantly improved through team decision making. This is our first main contribution to the flourishing literature on strategic sophistication. Another novel feature of our paper is our consideration of how first- and second-order beliefs are related. This yields insights into how individuals or teams perceive the consistency of their opponents. Similar to the findings for the relation between choices and first-order beliefs, we find also a difference between individuals and teams with respect to how consistent are first- and second-order beliefs. Teams expect their opponents – irrespective of whether this is an individual or a team – to play and expect the Nash strategy more often than individuals expect that from their opponents. Obviously, this insight contributes to a deeper understanding of why individuals and teams differ already with respect to how own choices relate to first-order beliefs. If a decision maker expects his opponent to be someone who is likely to play optimally by best responding to his beliefs, and in particular to play the Nash equilibrium and expect it from the other player, then such a decision maker has a stronger incentive to play Nash (and expect it from the opponent) than a decision maker who perceives the opponent as less consistent.

Our paper also adds to the literature on strategic sophistication through our analysis of the determinants of "Nash-consistency". This term refers to the situation where a decision maker plays Nash and expects the opponent to play Nash. While our analysis (in section 3.5.) has shown that team decision making – in situations where a dominant strategy is available – makes Nash-consistency more likely, our analysis also sheds light on the surprisingly low level of consistency in the first place. Like in our experiment with individuals, Costa-Gomes and Weizsäcker (2008) find consistent choices in only about 50% of cases. As one explanation for the high level of inconsistency between choices and first-order beliefs, Costa-Gomes and Weizsäcker (2008) show that subjects perceive a game differently when making choices than when stating beliefs (with stated beliefs revealing deeper strategic thinking than subjects' actions). We add through our analysis that consistency also depends on distributional preferences. If payoffs in the equilibrium are too much different or if a decision maker earns much less than the opponent in equilibrium then consistency between choices and first-order beliefs becomes less likely. It is easily conceivable, for instance, that inequality averse subjects might expect their opponent to play the equilibrium strategy (which can be a rational expectation), but nevertheless refrain from their equilibrium choice for reasons of payoff asymmetry. Our analysis has also revealed an influence of what we have called the risk of not picking the equilibrium choice. If there is more money at risk when a decision maker tries to reach his maximal possible payoff instead of playing Nash, then Nash-consistency becomes more likely to be observed. It seems important to note that the effects of distributional preferences and the money at risk when deviating from Nash are important both for individual and team decision making. These findings add to the explanation of Costa-Gomes and Weizsäcker (2008) for the low levels of consistency of choices and first-order beliefs.

Given that our paper addresses differences between individuals and teams it is also important to consider this paper's contribution to the literature on team decision making. First of all, it offers a classification of teams according to different strategic and non-strategic types of decision making. While the existing literature on team decision making has mainly focused on whether or not teams play equilibrium strategies more often than individuals (see, e.g., Bornstein et al., 2004; Cooper and Kagel, 2005), our experiment can also answer this question (affirmatively), but goes one step further by showing the distributional differences between individuals and teams with respect to eight different types of strategic and non-strategic behavior. In particular, the main difference (with respect to strategic types) originates from teams being more likely a *D1*-type, as explained above. This is an important difference as it shows that the higher strategic sophistication of teams is particularly driven by teams thinking one step further than individuals. This is in line with the findings of Kocher and Sutter (2005) where they have found in experimental guessing games that the modal depths of reasoning of teams is one level higher than the one of individuals.[14] Second, we do not only elicit actual choices, but also first- and second-order beliefs. This has not been done in the team decision making literature before. This approach allows checking the consistency of choices and first-order beliefs, respectively the expected consistency of one's opponent (by relating first- and second-order beliefs). In fact, we have found that teams are more consistent in this respect, and that they also expect their opponents to be more consistent than individuals expect their opponents to be.

Previous studies on team decision making can shed light on possible reasons for our findings of higher strategic sophistication of teams. Cooper and Kagel (2005), for instance, have shown that putting oneself into the shoes of the opponent in a signaling game increases

[14] On the limited depths of reasoning (of individuals) in guessing games, information cascade experiments or auctions see also Costa-Gomes and Crawford (2006), Kübler and Weizsäcker (2004), and Gneezy (2005), respectively.

the likelihood of strategic play (being defined in their game as the relative frequency with which an incumbent monopolist chooses a larger quantity than a myopic best reply to the threat of market entry by an opponent would require). Teams who talk about the perspective of the opponent are more likely to play strategically. It seems plausible that our teams have also been influenced by such thoughts. Likewise, Bosman, Hennig-Schmidt and van Winden (2006) have been seminal in analyzing the contents of communication in teams by videotaping them. They have found that teams discuss issues of fairness very prominently, and in a pretty much self-centered interpretation. This might also have happened in our teams, and it might explain why the variables intended to capture the fairness of payoffs in the Nash equilibrium have a significant influence on the likelihood of observing best replies of choices to own first-order beliefs. In our experiment, we have chosen not to observe team discussions in order to keep the environment as similar as possible to individual decision making. This advantage comes at the cost of being unable to analyze team discussions – which has been done in earlier studies from which we can nevertheless derive informed guesses about the driving forces of team decision making. However one might weigh the costs and benefits of our approach, we consider the findings reported in this experiment to provide valuable insights for two strands of literature, namely on strategic sophistication on the one-hand side and on team decision making on the other-hand side.

References

Bhatt, M., Camerer, C. F. (2005), Self-referential thinking and equilibrium as staes of mind in games: fMRI evidence. *Games and Economic Behavior* 52: 424-459.

Bornstein, G., Kugler, T., Ziegelmeyer, A. (2004). Individual and groups decisions in the centipede game: are groups more rational players? *Journal of Experimental Social Psychology* 40: 599-605.

Bornstein, G., Yaniv, I. (1998), Individual and group behavior in the ultimatum game: are groups more "rational" players? *Experimental Economics* 1: 101-108.

Bosman, R., Hennig-Schmidt, H., van Winden, F. (2006), Exploring group decision making in a power-to-take experiment. *Experimental Economics* 9: 35-51.

Cason, T., Mui, V.-L. (1997), A laboratory study of group polarization in the team dictator game. *Economic Journal* 107: 1465-1483.

Charness, G., Jackson, M. (2007), Group play in games and the role of consent in network formation. *Journal of Economic Theory* 136: 417-445.

Cooper, D. J., Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. *American Economic Review* 95: 477-509.

Costa-Gomes, M., Crawford, V. (2006), Cognition and behavior in guessing games: An experimental study. *American Economic Review* 96: 1737-1768.

Costa-Gomes, M., Crawford, V., Broseta, B. (2001), Cognition and behavior in normal-form games: An experimental study. *Econometrica* 69: 1193-1235.

Costa-Gomes, M., Weizsäcker, G. (2008), Stated beliefs and play in normal-form games. *The Review of Economic Studies* 75: 729-762.

Fehr, D., Kübler, D., Danz, D. (2008), Information and beliefs in a repeated normal-form game. Working Paper, Technical University of Berlin.

Feri, F., Irlenbusch, B., Sutter, M. (2009), Efficiency gains from team-based coordination – Large-scale experimental evidence. *American Economic Review*, forthcoming.

Fischbacher, U. (2007), z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics* 10: 171-178.

Gneezy, U. (2005), Step-level reasoning and bidding in auctions. *Management Science* 51: 1633-1642.

Greiner, B. (2004), An Online Recruiting System for Economic Experiments. In: Kremer, K., Macho, V. (eds.), Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Goettingen, Gesellschaft fuer wissenschaftliche Datenverarbeitung: 79-93.

Haruvy, E. (2002), Identification and testing of modes in beliefs. *Journal of Mathematical Psychology* 46: 88-109.

Haruvy, E., Stahl, D., Wilson, P. (1999), Evidence for optimistic and pessimistic behavior in normal-form games. *Economic Letters* 63: 255-259.

Kocher, M., Sutter, M. (2005), The decision maker matters: Individual versus group behaviour in experimental beauty contest games. *The Economic Journal* 115: 200-223.

Kübler, D., Weizsäcker, G. (2004), Limited depths of reasoning and failure of cascade formation in the laboratory. *Review of Economic Studies* 71: 425-441.

Rey-Biel, P. (2009), Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior* 65: 572-585.

Stahl, D., Wilson, P. (1994), Experimental evidence of players´ models of other players. *Journal of Economic Behavior and Organziation* 25: 309-327.

Stahl, D., Wilson, P. (1995), On players´ models of other players: theory and experimental evidence. *Games and Economic Behavior* 10: 213-254.

Weizsäcker, G. (2003), Ignoring the rationality of others: evidence from experimental normal-form games. *Games and Economic Behavior* 44: 145-171.

Tables and Figures

Table 1: Choices and beliefs of individuals and teams (relative frequencies in %)

		All games	*D-games*	*ND-games*
Choice	Nash Individuals	40.97 **	57.71 **	20.05
	Nash Teams	49.30	70.42	22.92
	Pareto Individuals	53.47 **	41.25 **	68.75
	Pareto Teams	45.72	28.96	66.67
	Other Individuals	5.56	1.04	11.20
	Other Teams	4.98	0.62	10.41
First-order beliefs	Nash Individuals	37.26	55.63	14.32
	Nash Teams	40.04	60.62	14.32
	Pareto Individuals	57.06	43.33	74.22
	Pareto Teams	55.21	38.75	75.78
	Other Individuals	5.68	1.04	11.46
	Other Teams	4.75	0.63	9.90
Second-order beliefs	Nash Individuals	30.55	47.71 **	9.11
	Nash Teams	33.79	53.96	8.59
	Pareto Individuals	62.62	50.21 *	78.13
	Pareto Teams	61.11	45.42	80.73
	Other Individuals	6.83	2.08 *	12.76
	Other Teams	5.10	0.62	10.68

** (*) significant difference between individuals (in a given row) and teams (in the next row) at $p < 0.05$ ($p < 0.1$); two-sided Mann-Whitney U-test.

Table 2: Complexity of the game, choices and beliefs (relative frequencies in %)

	Complexity (game type)[a]	*1R (D)*	*2R (D)*	*2R (ND)*	*3R (ND)*	*∞R (ND)*
Choice	Nash Individuals	76.67	38.75 **	31.25 *	14.58	17.19
	Nash Teams	81.67	59.17	44.79	10.42	18.23
	Pareto Individuals	22.50	60.00 **	68.75 *	56.25	75.00
	Pareto Teams	17.91	40.00	55.21	57.29	77.08
	Other Individuals	0.83	1.25	0.00	29.17	7.81
	Other Teams	0.42	0.83	0.00	32.29	4.69
First order beliefs	Nash Individuals	88.75	22.50	21.87 *	7.29	14.06
	Nash Teams	87.92	33.33	33.33	6.25	8.86
	Pareto Individuals	10.42	76.25	78.13 *	56.25	81.25 *
	Pareto Teams	10.83	66.67	66.67	57.29	89.58
	Other Individuals	0.83 *	1.25 *	0.00	36.46	4.69
	Other Teams	1.25	0.00	0.00	36.46	1.56
Second order beliefs	Nash Individuals	81.67	13.75 **	9.37	9.37	8.85
	Nash Teams	82.50	25.42	17.71	4.17	6.25
	Pareto Individuals	16.25	84.17 **	90.63 *	50.00	85.94
	Pareto Teams	17.08	73.75	82.29	57.29	91.67
	Other Individuals	2.08 *	2.08 *	0.00	40.63	5.21
	Other Teams	0.42 *	0.83	0.00	38.54	2.08

** (*) significant difference between individuals and teams at $p < 0.05$ ($p < 0.1$); two-sided Mann-Whitney U-test.

[a] The columns separate behavior according to (i) the different number of rounds (***R***) of iterated pure-strategy dominance a player needs to identify the own equilibrium choice, and (ii) the presence (***D***) or absence (***ND***) of a dominant strategy in the game.

Table 3: Consistency of decisions (relative frequency of best reply)

		All games	*D-games*	*ND-games*
Choice is best reply to first-order belief	consistent individuals	55.79**	59.79 **	50.78
	consistent teams	65.86	75.21	54.17
First-order belief is best reply to second-order belief	consistent individuals	59.26	61.67	56.25
	consistent teams	62.73	67.08	57.29

** (*) significant difference between individuals and teams at $p < 0.05$ ($p < 0.1$); two-sided Mann-Whitney U-test.

Table 4: Relative frequency of consistency-types Nash-CON and Max-CON

		all games	*DOM-games*	*WDOM-games*
	Nash-CON individuals	18.75**	27.71**	7.55
Player's own consistency	Nash-CON teams	27.66	45.00	5.99
	Max-CON individuals	32.41	31.46	33.59
	Max-CON teams	31.37	29.58	33.59
	Nash-CON individuals	10.53**	15.42**	4.43
Expected consistency of opponent	Nash-CON teams	17.36	28.75	3.13
	Max-CON individuals	44.56*	44.79**	44.27
	Max-CON teams	40.16	37.71	43.23

** (*) significant difference between individuals and teams at $p < 0.05$ ($p < 0.1$); two-sided Mann-Whitney U-test.

Table 5: Determinants of a player's consistency-type Nash-CON

[A] Probit regression

Variable	*Coefficient*	*Std. error*	*p-value*
dominant strategy	1.306	0.284	0.000
opponent dominant strategy	0.951	0.163	0.000
team player	-0.146	0.286	0.610
opponent = team	-0.057	0.241	0.813
team * opponent = team	0.056	0.329	0.863
team * dominant strategy	0.505	0.209	0.016
team * opponent dominant strategy	0.718	0.220	0.001
risk with maxstrategy	0.011	0.006	0.081
inequality in Nash	-0.026	0.012	0.028
difference to opponent's payoff in Nash	0.012	0.003	0.000
Constant	-1.316	0.383	0.001

N = 1.728; standard errors clustered for the 96 decision makers (48 individuals, 48 teams).

[B] Marginal effects of team decision making contingent on binary independent variables

Dominant strategy	Opponent dominant strategy	Opponent = Team	***Marginal effect***	***Std. error***	***p-value***
Yes	No	Yes	0.163	0.100	0.105
Yes	No	No	0.141	0.102	0.167
No	Yes	Yes	0.229	0.089	0.010
No	Yes	No	0.211	0.090	0.020
No	No	Yes	-0.008	0.025	0.737
No	No	No	-0.014	-0.029	0.624

Marginal effects evaluated at the mean of independent variables (risk with maxstrategy, inequality in Nash, difference to opponent's payoff in Nash) and contingent on the three binary variables indicated in the top row (and their interaction effect with team decision making).

Table 6: Determinants of the opponent's expected consistency-type being Nash-CON

[A] Probit regression

Variable	*Coefficient*	*Std. error*	*p-value*
dominant strategy	0.744	0.288	0.010
opponent dominant strategy	1.063	0.261	0.000
team player	-0.186	0.365	0.611
opponent =team	-0.011	0.257	0.965
team * opponent = team	0.035	0.341	0.918
team * dominant Strategy	0.705	0.229	0.002
team * opponent dominant strategy	0.565	0.269	0.036
risk with maxstrategy	0.005	0.006	0.435
inequality in Nash	-0.039	0.013	0.004
difference to opponent's payoff in Nash	-0.002	0.004	0.579
Constant	-1.197	0.346	0.001

N = 1.728; standard errors clustered for the 96 decision makers (48 individuals, 48 teams).

[B] Marginal effects of team decision making contingent on binary independent variables

Dominant strategy	Opponent dominant strategy	Opponent = Team	***Marginal effect***	*Std. error*	*p-value*
Yes	No	Yes	0.156	0.062	0.012
Yes	No	No	0.146	0.077	0.060
No	Yes	Yes	0.139	0.080	0.084
No	Yes	No	0.127	0.093	0.171
No	No	Yes	-0.009	0.018	0.598
No	No	No	-0.011	0.025	0.643

Marginal effects evaluated at the mean of independent variables (risk with maxstrategy, inequality in Nash, difference to opponent's payoff in Nash) and contingent on the three binary variables indicated in the top row (and their interaction effect with team decision making).

Table 7: Estimated types of individuals and teams – Own choices

	Individuals		Teams	
Type	p_k	ε_k	p_k	ε_k
Altruistic	.122	.232	.121	.445
	(.051)**	(.073)***	(.066)*	(.158)***
Pessimistic	.070	.553	.005	.999
	(.047)	(.128)***	(.212)	(.000)***
Naïve	.407	.220	.252	.208
	(.082)***	(.044)***	(.067)***	(.048)***
Equilibrium	.021	.097	.020	.318
	(.021)	(.094)	(.028)	(.171)*
Sophisticated	.058	.311	.020	.000
	(.051)	(.117)***	(.021)	(.000)
D1	.223	.504	.472	.315
	(.080)***	(.065)***	(.087)***	(.038)***
D2	.062	.242	.002	.999
	(.042)	(.101)**	(.201)	(.001)***
L2	.038	.376	.108	.270
	(.049)	(.147)**	(.060)*	(.076)***
Strategic	.402		.622	
	(.076)***		(.221)***	

(standard errors in parenthesis, *** significant at 1% level, ** significant at 5% level, * significant at 10% level)

Figure 1. The 18 normal-form games

game # 3 (*D*) [1R, 2R]

	1	2
1	72, 93	31, 46
2	84, 52	55, 79

game # 13 (*D*) [1R, 2R]

	1	2
1	94, 23	38, 57
2	45, 89	14, 18

game # 1 (*D*) [2R, 1R]

	1	2
1	75, 51	42, 27
2	48, 80	89, 68

game # 12 (*D*) [2R, 1R]

	1	2
1	21, 92	87, 43
2	55, 36	16, 12

game # 7 (*D*) [2R, 1R]

	1	2	3
1	59, 58	46, 83	85, 61
2	38, 29	70, 52	37, 23

game # 11 (*D*) [2R, 1R]

	1	2
1	31, 32	68, 46
2	72, 43	47, 61
3	91, 65	43, 84

game # 9 (*D*) [1R, 2R]

	1	2
1	28, 37	57, 58
2	22, 36	60, 84
3	51, 69	82, 45

game # 16 (*D*) [1R, 2R]

	1	2	3
1	42, 64	57, 43	80, 39
2	28, 27	39, 68	61, 87

game # 6 (*ND*) [3R, 2R]

	1	2
1	53, 86	24, 19
2	79, 57	42, 73
3	28, 23	71, 50

game # 15 (*ND*) [3R, 2R]

	1	2
1	76, 93	25, 12
2	43, 40	74, 62
3	94, 16	59, 37

game # 14 (*ND*) [2R, 3R)

	1	2	3
1	21, 26	52, 73	75, 44
2	88, 55	25, 30	59, 81

game # 2 (*ND*) [2R, 3R]

	1	2	3
1	42, 45	95, 78	18, 96
2	64, 76	14, 27	39, 61

game # 8 (*ND*) [∞R, ∞R]

	1	2	3
1	87, 32	18, 37	63, 76
2	65, 89	96, 63	24, 30

game # 10 (*ND*) [∞R, ∞R]

	1	2	3
1	67, 91	95, 64	31, 35
2	89, 49	23, 53	56, 78

game # 5 (*ND*) [∞R, ∞R]

	1	2
1	72, 59	26, 20
2	33, 14	59, 92
3	28, 83	85, 61

game # 4 (*ND*) [∞R, ∞R]

	1	2
1	46, 16	57, 88
2	71, 49	28, 24
3	42, 82	84, 60

game # 17 (*D*) [1R, 2R]

	1	2
1	22, 14	57, 55
2	30, 42	28, 37
3	15, 60	61, 88
4	45, 66	82, 31

game# 18 (*D*) [2R, 1R]

	1	2	3	4
1	56, 58	38, 29	89, 62	32, 86
2	15, 23	43, 31	61, 16	67, 46

Supplementary material (not intended for publication)

Table S1.Relative frequencies of choices in each game – Individuals

game nr. 3		(D)
	(70.83%)	(29.17%)
(37.5%)	72, 93	31, 46
(62.5%)	84, 52	55, 79

game nr. 13		(D)
	(45.83%)	(54.17%)
(87.5%)	94, 23	38, 57
(12.5%)	45, 89	14, 18

game nr. 1		(D)
	(75%)	(25%)
(37.5%)	75, 51	42, 27
(62.5%)	48, 80	89, 68

game nr. 12		(D)
	(87.5%)	(12.5%)
(62.5%)	21, 92	87, 43
(37.5%)	55, 36	16, 12

game nr. 7			(D)
	(0%)	(91.67%)	(8.33%)
(83.33%)	59, 58	46, 83	85, 61
(16.67%)	38, 29	70, 52	37, 23

game nr. 11		(D)
	(16.67%)	(83.33%)
(29.17%)	31, 32	68, 46
(8.33%)	72, 43	47, 61
(62.5%)	91, 65	43, 84

game nr. 9		(D)
	(37.5%)	(62.5%)
(0%)	28, 37	57, 58
(29.17%)	22, 36	60, 84
(70.83%)	51, 69	82, 45

game nr. 16			(D)
	(45.83%)	(4.17%)	(50%)
(66.67%)	42, 64	57, 43	80, 39
(33.33%)	28, 27	39, 68	61, 87

game nr. 6		(ND)
	(75%)	(25%)
(20.83%)	53, 86	24, 19
(79.17%)	79, 57	42, 73
(0%)	28, 23	71, 50

game nr. 15		(ND)
	(54.17%)	(45.83%)
(33.33%)	76, 93	25, 12
(12.5%)	43, 40	74, 62
(54.17%)	94, 16	59, 37

game nr. 14			(ND)
	(0%)	(29.17%)	(70.83%)
(25%)	21, 26	52, 73	75, 44
(75%)	88, 55	25, 30	59, 81

game nr. 2			(ND)
	(16.67%)	(41.67%)	(41.67%)
(70.83%)	42, 45	95, 78	18, 96
(29.17%)	64, 76	14, 27	39, 61

game nr. 8			(ND)
	(54.17%)	(12.5%)	(33.33%)
(25%)	87, 32	18, 37	63, 76
(75%)	65, 89	96, 63	24, 30

game nr. 10			(ND)
	(79.17%)	(8.33%)	(12.5%)
(83.33%)	67, 91	95, 64	31, 35
(16.67%)	89, 49	23, 53	56, 78

game nr. 5		(ND)
	(20.83%)	(79.17%)
(4.17%)	72, 59	26, 20
(25%)	33, 14	59, 92
(70.83%)	28, 83	85, 61

game nr. 4		(ND)
	(12.5%)	(87.5%)
(16.67%)	46, 16	57, 88
(12.5%)	71, 49	28, 24
(70.83%)	42, 82	84, 60

game nr. 17		(D)
	(58.33%)	(41.67%)
(0%)	22, 14	57, 55
(4.17%)	30, 42	28, 37
(29.17%)	15, 60	61, 88
(66.67%)	45, 66	82, 31

game nr. 18				(D)
	(0%)	(4.17%)	(20.83%)	(75%)
(41.67%)	56, 58	38, 29	89, 62	32, 86
(58.33%)	15, 23	43, 31	61, 16	67, 46

Table S2.Relative frequencies of choices in each game – Teams

game nr. 3		***(D)***
	(58.33%)	(41.67%)
(33.33%)	72, 93	31, 46
(66.67%)	84, 52	55, 79

game nr. 1		***(D)***
	(66.67%)	(33.33%)
(54.17%)	75, 51	42, 27
(45.83%)	48, 80	89, 68

game nr. 7			***(D)***
	(4.17%)	(83.33%)	(12.5%)
(54.17%)	59, 58	46, 83	85, 61
(45.83%)	38, 29	70, 52	37, 23

game nr. 9		***(D)***
	(45.83%)	(54.17%)
(0%)	28, 37	57, 58
(25%)	22, 36	60, 84
(75%)	51, 69	82, 45

game nr. 6		***(ND)***
	(58.33%)	(41.67%)
(0%)	53, 86	24, 19
(83.33%)	79, 57	42, 73
(16.67%)	28, 23	71, 50

game nr. 14			***(ND)***
	(0%)	(8.33%)	(91.67%)
(25%)	21, 26	52, 73	75, 44
(75%)	88, 55	25, 30	59, 81

game nr. 8			***(ND)***
	(75%)	(8.33%)	(16.67%)
(25%)	87, 32	18, 37	63, 76
(75%)	65, 89	96, 63	24, 30

game nr. 5		***(ND)***
	(20.83%)	(79.17%)
(16.67%)	72, 59	26, 20
(12.5%)	33, 14	59, 92
(70.83%)	28, 83	85, 61

game nr. 17		***(D)***
	(66.67%)	(33.33%)
(0%)	22, 14	57, 55
(0%)	30, 42	28, 37
(0%)	15, 60	61, 88
(100%)	45, 66	82, 31

game nr. 13		***(D)***
	(25%)	(75%)
(91.67%)	94, 23	38, 57
(8.33%)	45, 89	14, 18

game nr. 12		***(D)***
	(87.5%)	(12.5%)
(29.17%)	21, 92	87, 43
(70.83%)	55, 36	16, 12

game nr. 11		***(D)***
	(20.83%)	(79.17%)
(62.5%)	31, 32	68, 46
(4.17%)	72, 43	47, 61
(33.33%)	91, 65	43, 84

game nr. 16			***(D)***
	(70.83%)	(4.17%)	(25%)
(87.5%)	42, 64	57, 43	80, 39
(12.5%)	28, 27	39, 68	61, 87

game nr. 15		***(ND)***
	(29.17%)	(70.83%)
(25%)	76, 93	25, 12
(8.33%)	43, 40	74, 62
(66.67%)	94, 16	59, 37

game nr. 2			***(ND)***
	(8.33%)	(29.17%)	(62.5%)
(58.33%)	42, 45	95, 78	18, 96
(41.67%)	64, 76	14, 27	39, 61

game nr. 10			***(ND)***
	(79.17%)	(12.5%)	(8.33%)
(79.17%)	67, 91	95, 64	31, 35
(20.83%)	89, 49	23, 53	56, 78

game nr. 4		***(ND)***
	(25%)	(75%)
(4.17%)	46, 16	57, 88
(12.5%)	71, 49	28, 24
(83.33%)	42, 82	84, 60

game nr. 18			***(D)***	
	(0%)	(0%)	(20.83%)	(79.17%)
(41.67%)	56, 58	38, 29	89, 62	32, 86
(58.33%)	15, 23	43, 31	61, 16	67, 46

Table S3: Relative frequencies of first-order beliefs in each game – Individuals

game nr. 3 *(D)*

	(95.83%)	(4.17%)
(16.67%)	72, 93	31, 46
(83.33%)	84, 52	55, 79

game nr. 13 *(D)*

	(79.17%)	(20.83%)
(95.83%)	94, 23	38, 57
(4.17%)	45, 89	14, 18

game nr. 1 *(D)*

	(87.5%)	(12.5%)
(25%)	75, 51	42, 27
(75%)	48, 80	89, 68

game nr. 12 *(D)*

	(83.33%)	(16.67%)
(58.33%)	21, 92	87, 43
(41.67%)	55, 36	16, 12

game nr. 7 *(D)*

	(4.17%)	(91.67%)	(4.17%)
(54.17%)	59, 58	46, 83	85, 61
(45.83%)	38, 29	70, 52	37, 23

game nr. 11 *(D)*

	(8.33%)	(91.67%)
(20.83%)	31, 32	68, 46
(4.17%)	72, 43	47, 61
(75%)	91, 65	43, 84

game nr. 9 *(D)*

	(12.5%)	(87.5%)
(0%)	28, 37	57, 58
(8.33%)	22, 36	60, 84
(91.67%)	51, 69	82, 45

game nr. 16 *(D)*

	(8.33%)	(8.33%)	(83.33%)
(91.67%)	42, 64	57, 43	80, 39
(8.33%)	28, 27	39, 68	61, 87

game nr. 6 *(ND)*

	(83.33%)	(16.67%)
(4.17%)	53, 86	24, 19
(79.17%)	79, 57	42, 73
(16.67%)	28, 23	71, 50

game nr. 15 *(ND)*

	(83.33%)	(16.67%)
(20.83%)	76, 93	25, 12
(8.33%)	43, 40	74, 62
(70.83%)	94, 16	59, 37

game nr. 14 *(ND)*

	(8.33%)	(0%)	(91.67%)
(29.17%)	21, 26	52, 73	75, 44
(70.83%)	88, 55	25, 30	59, 81

game nr. 2 *(ND)*

	(4.17%)	(33.33%)	(62.5%)
(75%)	42, 45	95, 78	18, 96
(25%)	64, 76	14, 27	39, 61

game nr. 8 *(ND)*

	(83.33%)	(16.67%)	(0%)
(25%)	87, 32	18, 37	63, 76
(75%)	65, 89	96, 63	24, 30

game nr. 10 *(ND)*

	(87.5%)	(8.33%)	(4.17%)
(75%)	67, 91	95, 64	31, 35
(25%)	89, 49	23, 53	56, 78

game nr. 5 *(ND)*

	(12.5%)	(87.5%)
(20.83%)	72, 59	26, 20
(8.33%)	33, 14	59, 92
(70.83%)	28, 83	85, 61

game nr. 4 *(ND)*

	(16.67%)	(83.33%)
(4.17%)	46, 16	57, 88
(8.33%)	71, 49	28, 24
(87.5%)	42, 82	84, 60

game nr. 17 *(D)*

	(8.33%)	(91.67%)
(4.17%)	22, 14	57, 55
(0%)	30, 42	28, 37
(12.5%)	15, 60	61, 88
(83.33%)	45, 66	82, 31

game nr. 18 *(D)*

	(0%)	(0%)	(12.5%)	(87.5%)
(62.5%)	56, 58	38, 29	89, 62	32, 86
(37.5%)	15, 23	43, 31	61, 16	67, 46

Table S4: Relative frequencies of first-order beliefs in each game – Teams

game nr. 3		(D)
	(75%)	(25%)
(16.67%)	72, 93	31, 46
(83.33%)	84, 52	55, 79

game nr. 13		(D)
	(54.17%)	(45.83%)
(91.67%)	94, 23	38, 57
(8.33%)	45, 89	14, 18

game nr. 1		(D)
	(75%)	(25%)
(37.5%)	75, 51	42, 27
(62.5%)	48, 80	89, 68

game nr. 12		(D)
	(100%)	(0%)
(75%)	21, 92	87, 43
(25%)	55, 36	16, 12

game nr. 7			(D)
	(8.33%)	(83.33%)	(8.33%)
(70.83%)	59, 58	46, 83	85, 61
(29.17%)	38, 29	70, 52	37, 23

game nr. 11		(D)
	(4.17%)	(95.83%)
(20.83%)	31, 32	68, 46
(0%)	72, 43	47, 61
(79.17%)	91, 65	43, 84

game nr. 9		(D)
	(20.83%)	(79.17%)
(0%)	28, 37	57, 58
(16.67%)	22, 36	60, 84
(83.33%)	51, 69	82, 45

game nr. 16			(D)
	(50%)	(0%)	(50%)
(91.67%)	42, 64	57, 43	80, 39
(8.33%)	28, 27	39, 68	61, 87

game nr. 6		(ND)
	(54.17%)	(45.83%)
(0%)	53, 86	24, 19
(91.67%)	79, 57	42, 73
(8.33%)	28, 23	71, 50

game nr. 15		(ND)
	(66.67%)	(33.33%)
(12.5%)	76, 93	25, 12
(8.33%)	43, 40	74, 62
(79.17%)	94, 16	59, 37

game nr. 14			(ND)
	(0%)	(4.17%)	(95.83%)
(25%)	21, 26	52, 73	75, 44
(75%)	88, 55	25, 30	59, 81

game nr. 2			(ND)
	(4.17%)	(29.17%)	(66.67%)
(70.83%)	42, 45	95, 78	18, 96
(29.17%)	64, 76	14, 27	39, 61

game nr. 8			(ND)
	(95.83%)	(4.17%)	(0%)
(8.33%)	87, 32	18, 37	63, 76
(91.67%)	65, 89	96, 63	24, 30

game nr. 10			(ND)
	(100%)	(0%)	(0%)
(91.67%)	67, 91	95, 64	31, 35
(8.33%)	89, 49	23, 53	56, 78

game nr. 5		(ND)
	(16.67%)	(83.33%)
(12.5%)	72, 59	26, 20
(8.33%)	33, 14	59, 92
(79.17%)	28, 83	85, 61

game nr. 4		(ND)
	(25%)	(75%)
(0%)	46, 16	57, 88
(0%)	71, 49	28, 24
(100%)	42, 82	84, 60

game nr. 17		(D)
	(54.17%)	(45.83%)
(4.17%)	22, 14	57, 55
(0%)	30, 42	28, 37
(16.67%)	15, 60	61, 88
(79.17%)	45, 66	82, 31

game nr. 18			(D)	
	(0%)	(0%)	(12.5%)	(87.5%)
(62.5%)	56, 58	38, 29	89, 62	32, 86
(37.5%)	15, 23	43, 31	61, 16	67, 46

Table S5: Relative frequencies of second-order beliefs in each game – Individuals

	game nr. 3	***(D)***
	(87.5%)	(12.5%)
(20.83%)	72, 93	31, 46
(79.17%)	84, 52	55, 79

	game nr. 1	***(D)***
	(83.33%)	(16.67%)
(8.33%)	75, 51	42, 27
(91.67%)	48, 80	89, 68

	game nr. 7		***(D)***
	(8.33%)	(83.33%)	(8.33%)
(100%)	59, 58	46, 83	85, 61
(0%)	38, 29	70, 52	37, 23

	game nr. 9	***(D)***
	(12.5%)	(87.5%)
(0%)	28, 37	57, 58
(25%)	22, 36	60, 84
(75%)	51, 69	82, 45

	game nr. 6	***(ND)***
	(87.5%)	(12.5%)
(20.83%)	53, 86	24, 19
(79.17%)	79, 57	42, 73
(0%)	28, 23	71, 50

	game nr. 14		***(ND)***
	(8.33%)	(16.67%)	(75%)
(4.17%)	21, 26	52, 73	75, 44
(95.83%)	88, 55	25, 30	59, 81

	game nr. 8		***(ND)***
	(75%)	(8.33%)	(16.67%)
(0%)	87, 32	18, 37	63, 76
(100%)	65, 89	96, 63	24, 30

	game nr. 5	***(ND)***
	(8.33%)	(91.67%)
(4.17%)	72, 59	26, 20
(16.67%)	33, 14	59, 92
(79.17%)	28, 83	85, 61

	game nr. 17	***(D)***
	(25%)	(75%)
(4.17%)	22, 14	57, 55
(0%)	30, 42	28, 37
(25%)	15, 60	61, 88
(70.83%)	45, 66	82, 31

	game nr. 13	***(D)***
	(70.83%)	(29.17%)
(91.67%)	94, 23	38, 57
(8.33%)	45, 89	14, 18

	game nr. 12	***(D)***
	(87.5%)	(12.5%)
(75%)	21, 92	87, 43
(25%)	55, 36	16, 12

	game nr. 11	***(D)***
	(8.33%)	(91.67%)
(4.17%)	31, 32	68, 46
(4.17%)	72, 43	47, 61
(91.67%)	91, 65	43, 84

	game nr. 16		***(D)***
	(8.33%)	(16.67%)	(75%)
(79.17%)	42, 64	57, 43	80, 39
(20.83%)	28, 27	39, 68	61, 87

	game nr. 15	***(ND)***
	(83.33%)	(16.67%)
(25%)	76, 93	25, 12
(4.17%)	43, 40	74, 62
(70.83%)	94, 16	59, 37

	game nr. 2		***(ND)***
	(16.67%)	(20.83%)	(62.5%)
(95.83%)	42, 45	95, 78	18, 96
(4.17%)	64, 76	14, 27	39, 61

	game nr. 10		***(ND)***
	(79.17%)	(8.33%)	(12.5%)
(87.5%)	67, 91	95, 64	31, 35
(12.5%)	89, 49	23, 53	56, 78

	game nr. 4	***(ND)***
	(12.5%)	(87.5%)
(8.33%)	46, 16	57, 88
(4.17%)	71, 49	28, 24
(87.5%)	42, 82	84, 60

	game nr. 18		***(D)***	
	(8.33%)	(0%)	(16.67%)	(75%)
(87.5%)	56, 58	38, 29	89, 62	32, 86
(12.5%)	15, 23	43, 31	61, 16	67, 46

Table S6: Relative frequencies of second-order beliefs in each game – Teams

game nr. 3		***(D)***
	(91.67%)	(8.33%)
(25%)	72, 93	31, 46
(75%)	84, 52	55, 79

game nr. 1		***(D)***
	(83.33%)	(16.67%)
(8.33%)	75, 51	42, 27
(91.67%)	48, 80	89, 68

game nr. 7			***(D)***
	(4.17%)	(79.17%)	(16.67%)
(87.5%)	59, 58	46, 83	85, 61
(12.5%)	38, 29	70, 52	37, 23

game nr. 9		***(D)***
	(20.83%)	(79.17%)
(0%)	28, 37	57, 58
(29.17%)	22, 36	60, 84
(70.83%)	51, 69	82, 45

game nr. 6		***(ND)***
	(87.5%)	(12.5%)
(4.17%)	53, 86	24, 19
(87.5%)	79, 57	42, 73
(8.33%)	28, 23	71, 50

game nr. 14			***(ND)***
	(0%)	(4.17%)	(95.83%)
(12.5%)	21, 26	52, 73	75, 44
(87.5%)	88, 55	25, 30	59, 81

game nr. 8			***(ND)***
	(91.67%)	(0%)	(8.33%)
(8.33%)	87, 32	18, 37	63, 76
(91.67%)	65, 89	96, 63	24, 30

game nr. 5		***(ND)***
	(8.33%)	(91.67%)
(4.17%)	72, 59	26, 20
(4.17%)	33, 14	59, 92
(91.67%)	28, 83	85, 61

game nr. 17		***(D)***
	(20.83%)	(79.17%)
(0%)	22, 14	57, 55
(0%)	30, 42	28, 37
(4.17%)	15, 60	61, 88
(95.83%)	45, 66	82, 31

game nr. 13		***(D)***
	(62.5%)	(37.5%)
(95.83%)	94, 23	38, 57
(4.17%)	45, 89	14, 18

game nr. 12		***(D)***
	(87.5%)	(12.5%)
(45.83%)	21, 92	87, 43
(54.17%)	55, 36	16, 12

game nr. 11		***(D)***
	(20.83%)	(79.17%)
(25%)	31, 32	68, 46
(4.17%)	72, 43	47, 61
(70.83%)	91, 65	43, 84

game nr. 16			***(D)***
	(29.17%)	(4.17%)	(66.67%)
(87.5%)	42, 64	57, 43	80, 39
(12.5%)	28, 27	39, 68	61, 87

game nr. 15		***(ND)***
	(75%)	(25%)
(25%)	76, 93	25, 12
(0%)	43, 40	74, 62
(75%)	94, 16	59, 37

game nr. 2			***(ND)***
	(4.17%)	(20.83%)	(75%)
(79.17%)	42, 45	95, 78	18, 96
(20.83%)	64, 76	14, 27	39, 61

game nr. 10			***(ND)***
	(91.67%)	(8.33%)	(0%)
(87.5%)	67, 91	95, 64	31, 35
(12.5%)	89, 49	23, 53	56, 78

game nr. 4		***(ND)***
	(4.17%)	(95.83%)
(4.17%)	46, 16	57, 88
(4.17%)	71, 49	28, 24
(91.67%)	42, 82	84, 60

game nr. 18			***(D)***	
	(0%)	(0%)	(29.17%)	(70.83%)
(62.5%)	56, 58	38, 29	89, 62	32, 86
(37.5%)	15, 23	43, 31	61, 16	67, 46

Table S7: Relative frequency of correct higher order beliefs (%) – Individuals (row players)

	First – order beliefs		***Second – order beliefs***	
game	***correct***	***incorrect***	***correct***	***incorrect***
game #1	62.5	37.5	62.50	37.50
game #2	54.17	45.83	62.50	37.50
game #3	70.83	29.17	75.00	25.00
game #4	62.5	37.5	75.00	25.00
game #5	83.33	16.67	62.50	37.50
game #6	70.83	29.17	66.67	33.33
game #7	70.83	29.17	58.33	41.67
game #8	41.67	58.33	62.50	37.50
game #9	62.50	37.50	70.83	29.17
game #10	75.00	25.00	70.83	29.17
game #11	66.67	33.33	66.67	33.33
game #12	79.17	20.83	50.00	50.00
game #13	45.83	54.17	87.50	12.50
game #14	70.83	29.17	70.83	29.17
game #15	50.00	50.00	62.50	37.50
game #16	37.50	62.50	62.50	37.50
game #17	50.00	50.00	75.00	25.00
game #18	54.17	45.83	58.33	41.67
all games	***61.57***	***38.43***	***66.67***	***33.33***

Table S8: Relative frequency of correct higher order beliefs (%) – Individuals (column players)

	First – order beliefs		***Second – order beliefs***	
game	***correct***	***incorrect***	***correct***	***incorrect***
game #1	58.33	41.67	70.83	29.17
game #2	54.17	45.83	37.50	62.50
game #3	58.33	41.67	79.17	20.83
game #4	66.67	33.33	62.50	37.50
game #5	50.00	50.00	83.33	16.67
game #6	66.67	33.33	83.33	16.67
game #7	45.83	54.17	66.67	33.33
game #8	66.67	33.33	75.00	25.00
game #9	62.50	37.50	75.00	25.00
game #10	79.17	20.83	75.00	25.00
game #11	33.33	66.67	87.50	12.50
game #12	54.17	45.83	79.17	20.83
game #13	83.33	16.67	70.83	29.17
game #14	70.83	29.17	70.83	29.17
game #15	50.00	50.00	70.83	29.17
game #16	70.83	29.17	54.17	45.83
game #17	70.83	29.17	70.83	29.17
game #18	45.83	54.17	58.33	41.67
all games	***60.42***	***39.58***	***70.60***	***29.40***

Table S9: Relative frequency of correct higher order beliefs (%) – Teams (row players)

	First – order beliefs		***Second – order beliefs***	
game	***correct***	***incorrect***	***correct***	***incorrect***
game #1	66.67	33.33	70.83	29.17
game #2	62.50	37.50	66.67	33.33
game #3	45.83	54.17	66.67	33.33
game #4	79.17	20.83	83.33	16.67
game #5	62.50	37.50	70.83	29.17
game #6	62.50	37.50	75.00	25.00
game #7	58.33	41.67	75.00	25.00
game #8	62.50	37.50	87.50	12.50
game #9	45.83	54.17	58.33	41.67
game #10	79.17	20.83	83.33	16.67
game #11	83.33	16.67	62.50	37.50
game #12	70.83	29.17	58.33	41.67
game #13	41.67	58.33	87.50	12.50
game #14	79.17	20.83	66.67	33.33
game #15	37.50	62.50	58.33	41.67
game #16	66.67	33.33	70.83	29.17
game #17	62.50	37.50	83.33	16.67
game #18	70.83	29.17	58.33	41.67
all games	***63.19***	***36.81***	***71.30***	***28.70***

Table S10: Relative frequency of correct higher order beliefs (%) – Teams (column players)

	First – order beliefs		***Second – order beliefs***	
game	***correct***	***incorrect***	***correct***	***incorrect***
game #1	58.33	41.67	75.00	25.00
game #2	54.17	45.83	70.83	29.17
game #3	54.17	45.83	70.83	29.17
game #4	79.17	20.83	79.17	20.83
game #5	54.17	45.83	70.83	29.17
game #6	75.00	25.00	54.17	45.83
game #7	50.00	50.00	70.83	29.17
game #8	58.33	41.67	79.17	20.83
game #9	83.33	16.67	58.33	41.67
game #10	66.67	33.33	83.33	16.67
game #11	50.00	50.00	70.83	29.17
game #12	37.50	62.50	79.17	20.83
game #13	83.33	16.67	29.17	70.83
game #14	66.67	33.33	87.50	12.50
game #15	41.67	58.33	45.83	54.17
game #16	75.00	25.00	45.83	54.17
game #17	70.83	29.17	62.50	37.50
game #18	62.50	37.50	70.83	29.17
all games	***62.27***	***37.73***	***66.90***	***33.10***

Experimental instructions

In the following we present the instructions for the individual treatment as a baseline and indicate parts which were used for the team treatment by underlined letters and parts which were used for the mixed treatment by italic letters.

18 games in which two randomly chosen cabins are interacting with each other

In this experiment you will make decisions in 18 different “matrix-games”. In each „matrix-game“ **two different cabins** are **interacting** with each other. From a total of eight cabins, in every game two cabins are randomly selected to interact with each other. *Out of those two cabins there is always one cabin with three participants and one cabin with one participant. This means in this experiment interaction takes place between teams of three participants and single players. It is randomly selected whether you are a member of a team or a single player by drawing cards at the beginning of the experiment.*

In each “matrix game” both cabins have **up to four choices**. In most games the quantity of choices is two or three. All choices in your own cabin are indicated with the following signs:

- §
- %
- @
- µ

Choices in the other cabin are indicated with the following signs:

- §§
- %%
- @@
- µµ

The combination of choices of both randomly selected cabins decides about your potential payment in the experiment.

Three types of screens in each game (three decisions in each game)

For each of the 18 games you will be randomly matched with another cabin. Within each game you need to make three different choices on three different screen types. For those three decisions the matching of the cabins remains constant.

Please take care of the fact that the sequence of those three decisions is randomly determined. Because of this you always need to pay attention to the upper row placed above the matrices. On the following three pages you will find examples of the three different screen-types (decisions). Note that the numbers within the matrices are just examples.

Payment

During the experiment you will have to make 18 * 3 decisions. At the end of the experiment one of the 18 games and one of the corresponding three decisions will be randomly assigned to be relevant for your payment.

This random selection is done by a draw from a deck of cards showing numbers from 1 to 18 (to select one of the 18 games) and from a deck of cards showing numbers 1 to 3 (to select one of the three decisions). You will draw the card by **yourself**. In the following you will find an explanation how much you can earn.

Independently from your decisions you will receive a show up fee of **4 Euros**.

Feedback

Until the end of period 18 you will not get any feedback about the choices of the other cabins. As soon as the game and the screen-type (decision) which will be relevant for payment are selected, you will get detailed information about the corresponding decisions and the payment resulting from these decisions.

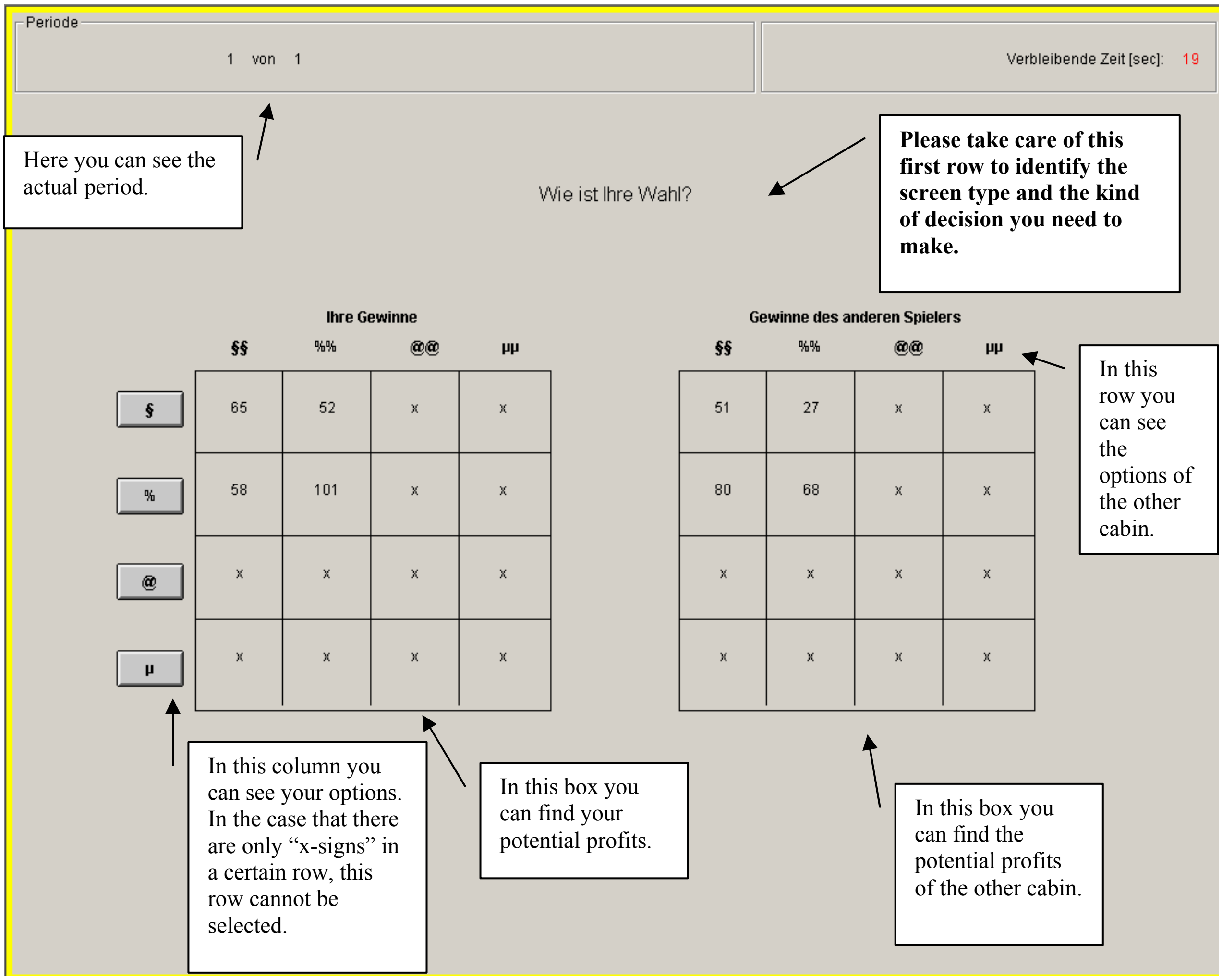
Teams in cabins

*In the case that there are three participants in your cabin you need to agree on **one common team decision**. At the end of the experiment **each team member gets** the **payment as explained** in the instructions.*

Teams in cabins

In this experiment there is a team consisting of three members in each cabin. The team members are required to agree on **one common team decision**. At the end of the experiment **each team member gets** the **payment as explained** in the instructions.

Screen type A. What is your choice?



On the left hand side of the screen you can find your potential profit, depending on your own choice (in this example „§" oder „%") and on the choice of the cabin that is matched with you in a certain period (in this example „§§" oder „%%").
Lets assume that you choose the option „%", while the other cabin chooses the option „§§", in this particular case you would earn 58 experimental points while the other cabin earns 80 experimental points.
(Please take care that the screen for the other cabin is transposed (looks contrariwise). If you have any questions about this, please ask the instructor of the experiment)

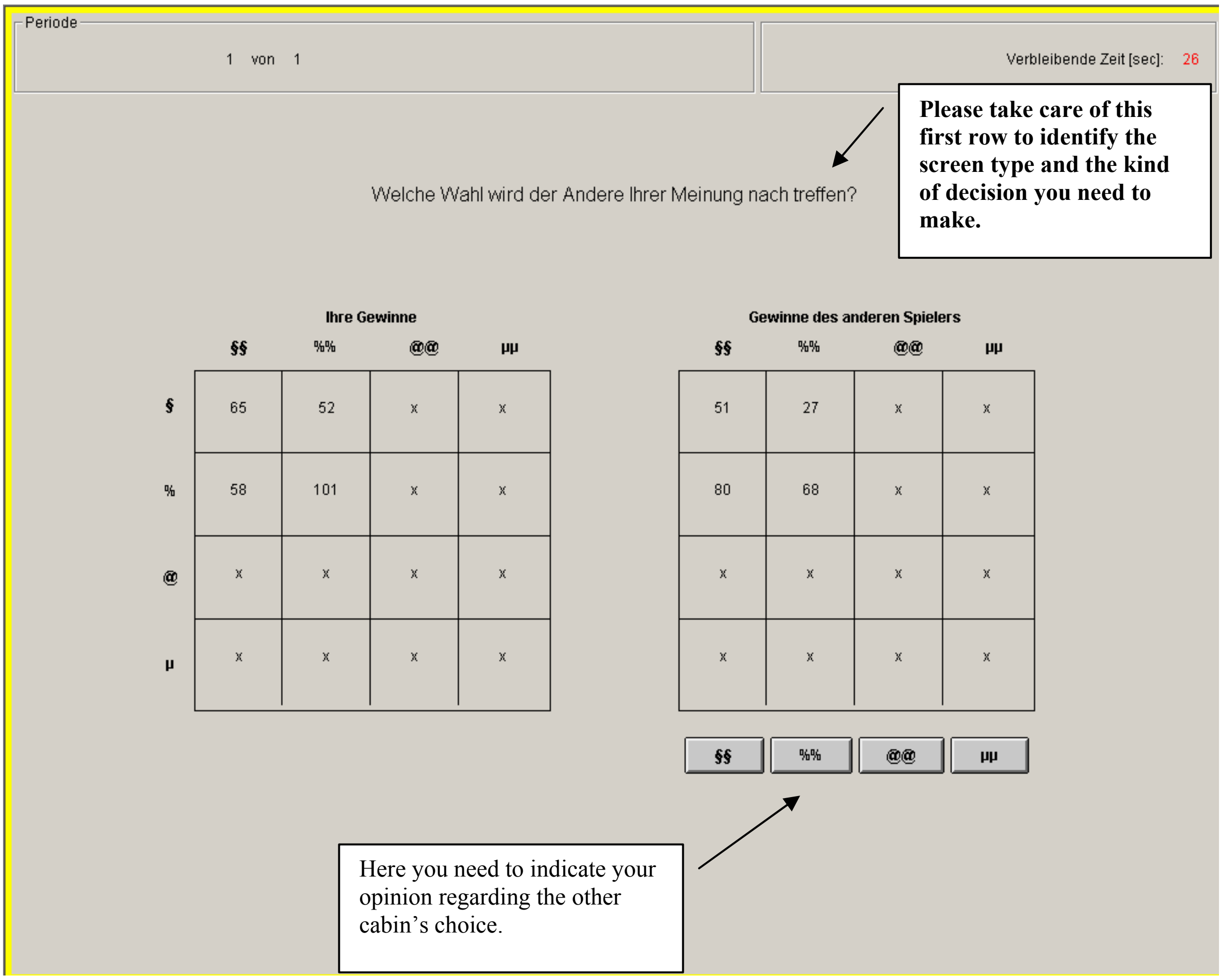
In the case that at the end of the experiment screen type A is randomly selected to be relevant for your payment, your profit (measured in experimental points) will be determined by the combination of your own choice and the choice of the other cabin.

The relevant exchange rate between experimental points and Euros is:

1 experimental point = 0.30 Euros

In the example illustrated above this would result in a payoff of 14.4 Euros (plus show up fee) for yourself and 24 Euros (plus show up fee) for the other cabin.

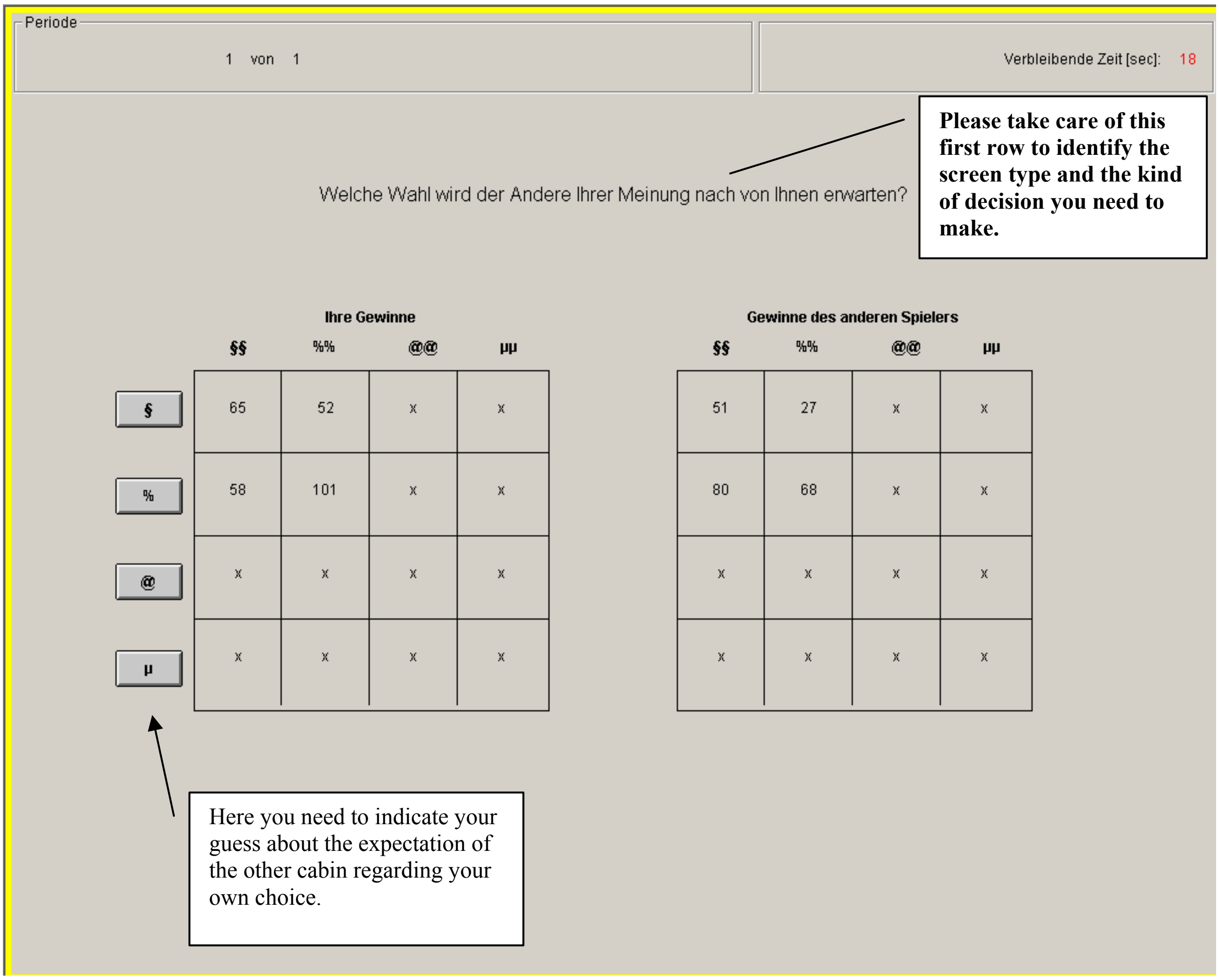
Screen type B. According to your opinion, which option will be chosen by the other cabin?



In the case of screen type B you need to correctly predict the decision of the other cabin. In order to do that you need to select the choice which you expect from the other cabin in the screen-area as explained above.

In the case that at the end of the experiment screen type B is randomly selected to be relevant for your payment, your profit will be **15 Euros** if your guess was **correct**. To assess whether your guess was correct we will compare your expectation indicated on this screen with the real choice of the other cabin.

Screen type C. What is the choice that the other player expects from you?



In the case of screen type C you need to correctly predict the other cabin's expectation regarding your own choice.
On screen type B (which is explained on the previous page) you need to correctly predict the decision of the other cabin. For each game the other cabin needs to make the same prediction which means that the other cabin needs to state an expectation regarding your own choice. On screen type C you need to indicate which choice the other cabin expects from you.

In the case that at the end of the experiment screen type C is randomly selected to be relevant for payment, your profit will be **15 Euros** if your guess regarding the expectation of the other cabin was **correct**.
To assess whether your guess was correct we will compare your decision on this screen (which is your guess regarding the expectation of the other cabin) with the real expectation of the other cabin.

Control questions

Please answer the following questions which are all related to the matrices of the three screen-types explained before.

Screen-type A.

1) Let's assume you choose option „§" and the other cabin chooses option „%%":
   a. How many experimental points do you earn? ______________
   b. How many experimental points does the other cabin earn? _______________
2) Let's assume you choose option „%" and the other cabin chooses option „§§":
   a. How many experimental points do you earn? ______________
   b. How many experimental points does the other cabin earn? _______________

Screen-type B.

3) Let's assume you expect the other cabin to choose option „%%" and the other cabin chooses option „§§".
   a. How much do you earn (show up fee excluded), if this screen-type (decision) is selected to be relevant for payment? _________________
4) Let's assume you expect the other cabin to choose option „§§" and the other cabin chooses option „§§".
   a. How much do you earn (show up fee excluded), if this screen-type (decision) is selected to be relevant for payment? _________________

Screen-type C.

5) Let's assume the other cabin expects you to choose option „%" and let's assume you select option „%" on screen-type C.
   a. How much do you earn (show up fee excluded), if this screen-type (decision) is selected to be relevant for payment? _________________
6) Let's assume the other cabin expects you to choose option „§"and let's assume you select option „%" on screen-type C.
   a. How much do you earn (show up fee excluded), if this screen-type (decision) is selected to be relevant for payment? _________________