What drives people's choices in turn-taking games, if not game-theoretic rationality?

Sujata Ghosh

Indian Statistical Institute
Chennai, India

sujata@isichennai.res.in

Aviad Heifetz

The Open University of Israel
Raanana, Israel

aviadhe@openu.ac.il

Rineke Verbrugge

University of Groningen
Groningen, The Netherlands

L.C.Verbrugge@rug.nl

Harmen de Weerd

University of Groningen
Groningen, The Netherlands

h.a.de.weerd@rug.nl

In an earlier experiment, participants played a perfect information game against a computer, which was programmed to deviate often from its backward induction strategy right at the beginning of the game. Participants knew that in each game, the computer was nevertheless optimizing against some belief about the participant's future strategy. In the aggregate, it appeared that participants applied forward induction. However, cardinal effects seemed to play a role as well: a number of participants might have been trying to maximize expected utility.

In order to find out how people really reason in such a game, we designed centipede-like turn-taking games with new payoff structures in order to make such cardinal effects less likely. We ran a new experiment with 50 participants, based on marble drop visualizations of these revised payoff structures. After participants played 48 test games, we asked a number of questions to gauge the participants' reasoning about their own and the opponent's strategy at all decision nodes of a sample game. We also checked how the verbalized strategies fit to the actual choices they made at all their decision points in the 48 test games.

Even though in the aggregate, participants in the new experiment still tend to slightly favor the forward induction choice at their first decision node, their verbalized strategies most often depend on their own attitudes towards risk and those they assign to the computer opponent, sometimes in addition to considerations about cooperativeness and competitiveness.

1 Introduction

In game theory, turn-taking games are represented by game trees referred to as extensive-form games. It is well known that the concept of *game-theoretic rationality* provides a canonical approach towards solving extensive-form games with perfect information, namely *Backward Induction* (BI), which yields a subgame-perfect equilibrium (which is a unique in generic turn-taking games with no pay-off ties) [40, 41]. However, BI is often criticized for not taking into account that a player may end up in one particular subgame rather than another subgame, e.g. after a deviation of other players from their BI strategy. Thus, the past moves and reasoning of the players are not taken into consideration, only the future. All players commonly believe in everybody's future rationality, no matter how irrational players' past behavior has already proven. In some cases, such complete lack of consideration of previous moves may be suboptimal for players.

An alternative approach, that of *Forward Induction* (FI), does take these previous moves of the opponent(s) under consideration and tries to rationalize the opponent's past behavior for assessing his future

moves. Thus, when a player is about to play in a subgame that has been reached due to some strategy of the opponent that is not consistent with both common knowledge (belief) of rationality for each of the players *and* his past behavior, the player may still rationalize the opponent's past behavior. She may attribute to her opponent a strategy that is optimal against a possible suboptimal strategy of hers, or a strategy that is optimal against some rational strategy of hers, which is only optimal against a suboptimal strategy of his, and so on. If the player pursues this kind of rationalizing reasoning to the highest extent possible and reacts accordingly, she ends up choosing what is called an Extensive-Form Rationalizable (EFR) strategy [5,6,39]. For perfect information games, to which we restrict our attention in this paper, see Definition 1 in [26], which we we will be using here.

Even though EFR strategies may be distinct from BI strategies (e.g., see [43], Game 1 in Figure 1), in perfect-information games without relevant payoff ties it has be been shown that the unique EFR outcome coincides with the unique BI outcome [5,6,26] . In case there are relevant payoff ties, however, EFR outcomes may form a—possibly strict—subset of the BI outcomes [10,11,26,41], see game 3 in Figure 1.

In [19], we asked the question: *Are people inclined to use forward induction when they play a game?* Our pivotal interest was to examine participants' behavior following a deviation from BI behavior by their opponent right at the beginning of the game. We designed a Marble Drop game experiment where the participants played several rounds of turn-taking games with the computer. Related research has been done on how people reason in dynamic games, also focusing on forward induction [4,8,12,27,46]. However, their games are not perfect information, in contrast to our Marble Drop games. In these games, we programmed the computer so as to follow, in each repetition of each game, a strategy which is optimal with respect to some strategy of the human participant. We provided a more intuitive framing of the game trees, inspired by Meijering et al.'s Marble Drop [32,33]. In [19], it turned out that in the aggregate, people's first decisions could be explained as EFR behavior. However, it also seemed that in many cases, cardinal effects could have played a role. Consequently, we wanted to design a new Marble Drop experiment which minimizes this cardinal effect, and in the process, improve certain other aspects of our previous experiment, based on our own findings and discussion with colleagues working in empirical research.

To summarize, the main differences between the current Marble Drop experiment and the one reported in [19] are as follows. The aim is to have better overall understanding of the participants' behavior:

- At the beginning of the experiment, we explained verbally in a more comprehensive way regarding how the computer agent was programmed (including "Imagine that you are playing against a different computer opponent each time" and "the computer thinks that you already have a plan for that game, and it plays the best response to the plan it thinks that you have for that game, however, the computer does not learn from previous games and does not take into account your choices during the previous games".)
- We changed the payoff structures of the experimental games in order to minimize cardinal effects on the behavior of the participants.
- We modified the questions asked to the Group A and Group B participants to direct some of the participants towards strategic reasoning while playing the games and note the effects of such prompts.
- We asked more structured and pointed questions about the participants' decision-making process at the end of the experiment to bring out the reasoning behind their behaviors in a more explicit manner.
-We changed the monetary incentive so that a wealth effect would be eradicated. By basing payment on number of marbles in a randomly drawn game, the participants perceived a clear difference between earning k and $k+1$ marbles in a game (not just 5 cents as in the previous experiment, but 3.75 euros).

We mainly focused on the following questions:

1) Do participants now more clearly play according to FI?
2) If not, what are they actually doing? What roles are played by risk attitudes and cooperativeness versus competitiveness?
3) Can they be reasonably divided into types of players?

Additionally, we are interested in whether people take the perspective of their opponent and make use of *theory of mind* while playing the Marble Drop experiment. Theory of mind refers to the ability to reason about unobservable mental content of others, such as beliefs, desires, or goals [42]. This theory of mind ability can even be used to reason about the theory of mind of others, and reason about the way others reason about beliefs and goals. This *second-order theory of mind* allows us to understand sentences such as "Alice *knows* that Bob *wants* to get six marbles", and use this to adjust our predictions of the behavior of Alice.

The rest of this paper is structured as follows. In Section 2, we explain the new experimental games and the differences with the previous experiment [19]. Section 3 presents the experimental results about the participants' decisions in the games as well as their reports about their own reasoning behind their behaviors. Section 4 discusses the results and provides suggestions for future research.

2   A Marble Drop Experiment

In [19], we designed a marble drop game experiment to investigate whether people are inclined to use Forward Induction (FI) when they play dynamic perfect-information games. The participants played 8 rounds of turn-based games against a computer opponent, repeating in each round a set of 6 games, distinct in terms of payoff structures. By letting participants play against a computer opponent, we ensure that each participant encounters the same situations, which allows us to eliminate *variability due to the strategy of the opponent* in our analysis. In these two-player games, the players played alternately. Let C denote the computer, and let P denote the participant. In four of the games the computer plays first, followed by the participant, and each of the players can play at two decision nodes. In the remaining two games, which are truncated versions of two of the games described earlier, the participant gets first chance to move. It appeared that participants did apply FI in playing these games, that is, in all likelihood the participants responded in a way which is optimal with respect to the conjecture that the computer is after a larger prize than the one it has foregone, even when this necessarily meant that the computer has attributed future irrationality to the participant when the computer made the first move in the games. However, a closer look at the individual choices revealed that cardinal effects probably played a role in these choices.

To get more conclusive evidence for FI reasoning on the part of the participants, we designed the current experiment, where the participants once again played 8 rounds of turn-based games, repeating in each round a set of 6 games as earlier. The main difference from the earlier experiment is in the payoff structures of the 6 games used for the experiment. We revised the payoff structures so as to minimize the probable cardinal effects reported in [19]. The main features of the new payoff structures (cf. Figures 1 and 2) are as follows:
- Game 1 differs from Game 2 only in the payoff of the Computer (C) following action a (ending the game right away). As a result, Game $1'$ is the truncation of Game 1 as well as Game 2. This means that participants who get to play in Game 1 and Game 2 face exactly the same continuation game (Game $1'$).
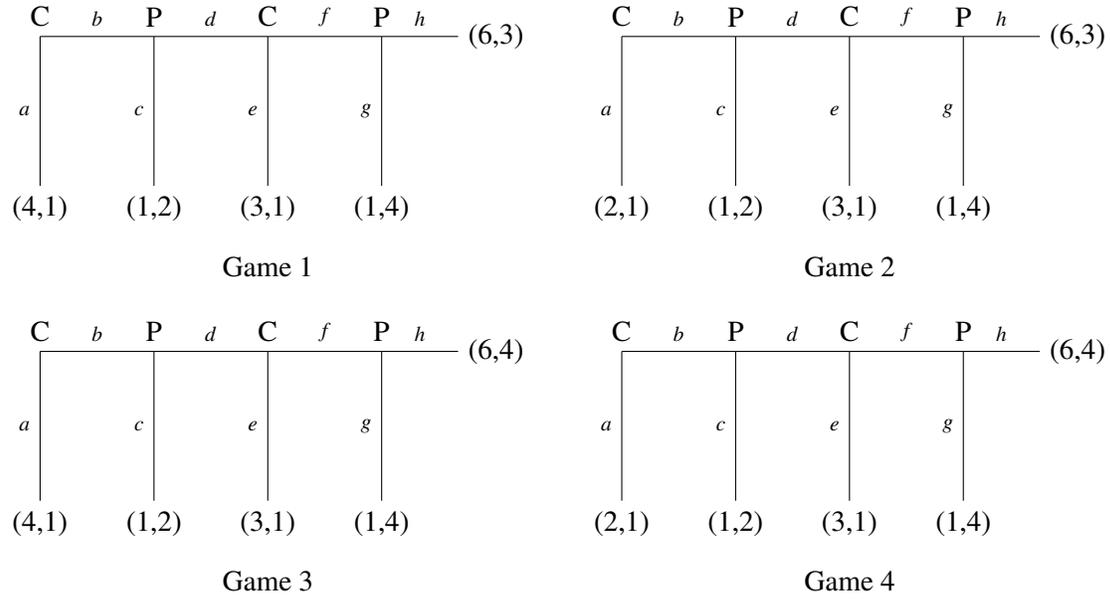
Figure 1: Collection of the main games used in the experiment. The ordered pairs at the leaves represent payoffs for the computer (C) and the participant (P), respectively.
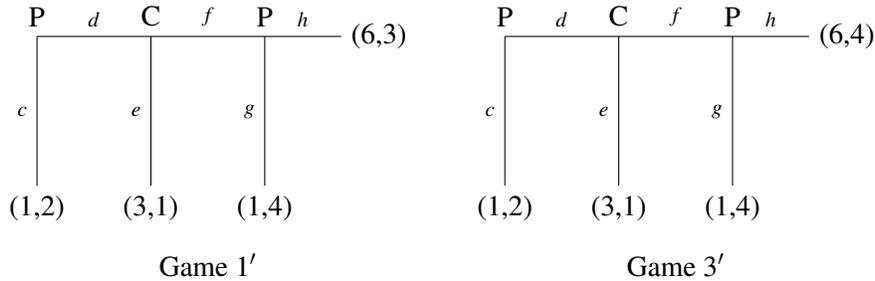

Figure 2: Truncated versions of Game 1 and Game 3.

We assumed that these new payoffs would enable an unobstructed comparison between the participants' behavior across these games (i.e. without confounding effects of cardinal differences in payoffs, which may have interfered in the previous experiment [19]). In [19], Game 1 and Game 2 had the same tree structures as currently, but the payoffs of player C following a and e were interchanged, so that Game 1′ was not a truncation of Game 2, hence some cardinal effect may have been instrumental in participants' choices at the nodes of these games.

- The same holds good for Game 3 and Game 4, respectively.

- Game 1 differs from Game 3 only in P's payoff following h at the very end. One can compare Game 2 and Game 4 in a similar manner. Previously, the payoff structure of Game 1 differed from that of Game 3, and the same for Game 2 and Game 4, respectively.

- We removed all the zero payoffs from these games, which some participants experienced as particularly "bad" in the previous experiment (according to their verbal reports).

- At C's second chance to move, exiting by playing e guarantees C a payoff of 3, which is only slightly

smaller than the expected payoff 3.5 of a fifty-fifty lottery between g (with which C gets 1) and h (with which C gets 6). This means that if at an earlier node, P believes that C is "confused" (because C has just deviated from its BI behavior), by attributing to C a fifty-fifty belief on P's future behavior at the last node (if reached), P cannot conclude what C will do next (choose e or f) if P believes that C is mildly risk-averse, because the precise level of C's presumed risk aversion would determine whether a sure payoff of 3 is preferable or not to a fifty-fifty lottery between 1 and 6 in C's eyes.

Furthermore, if due to the above, participant P has a 50-50 belief on C's future choice between e or f and P is herself mildly risk-averse, cardinal considerations on their own will not give her a clear guidance what to do: exiting by playing c will guarantee the payoff 2, while continuing by playing d yields a fifty-fifty lottery between 1 (if C chooses e) or 4 (if C chooses f), with a slightly higher expected payoff of 2.5.

Thus, participants who are unsure how to interpret C's initial deviation from BI will neither be able to come up with an easy forecast regarding C's future behavior on the basis of cardinal payoff considerations, nor will they have a strong preference for c or d on the basis of cardinal payoff considerations—which were a confounding potential explanation for participants' behavior in the previous experiment, which we would like to circumvent.

To get some idea about BI and FI strategies in the current experimental games, let us now concentrate on Game 1 (a variant of Reny's game [43]). (For the remaining games, the reader may adapt the reasoning strategies discussed in [19].) In Game 1, the unique Backward Induction (BI) strategy for player C is $a;e$, while for player P it is $c;g$. In case the last decision node of the game is reached, player P will play g (which will give P a better payoff at that node) yielding 1 for C. Thus, in the previous node, if reached, C will play e to be better off. Continuing like this from the end to the start of the game by BI reasoning, it can be inferred that whoever is the current player will choose to end the game immediately.

Forward induction, in contrast, would proceed as follows. Among the two strategies of player C that are compatible with reaching the first decision node of player P, namely $b;e$ and $b;f$, only the latter is rational for player C. This is because $b;e$ is dominated by $a;e$, while $b;f$ is optimal for player C if she believes that player P will play $d;h$ with a high enough probability. Attributing to player C the strategy $b;f$ is thus player P's best way to rationalize player C's choice of b, and in reply, $d;g$ is player P's best response to $b;f$. Thus, the unique Extensive-Form Rationalizable (EFR) strategy of P is $d;g$, which is distinct from her BI strategy $c;g$. Nevertheless, player C's best response to $d;g$ is $a;e$, which is therefore player C's EFR strategy. Hence the EFR outcome of the game (with the EFR strategies $a;e$ and $d;g$) is identical to the BI outcome. A summary of the strategies in the games of Figures 1 and 2 is given in Table 1.

We now discuss the experimental procedure, which is almost the same as reported in [19]. A group of 50 Bachelor's and Master's students from different disciplines participated in the experiment. As in [19], the participants had little or no knowledge of game theory, so as to ensure that neither backward induction nor forward induction reasoning was already known to them. 96% of the participants were aged between 18 and 32. 48% of the participants were female, 52% of the participants were male. All experiments were conducted at the Institute of Artificial Intelligence at the University of Groningen, in the ALICE Lab.

The participants played the finite perfect-information games in a graphical interface on the computer screen (cf. Figure 3). In each game, a marble was about to drop, and both the participant and the computer determined its path: The participant controlled the orange trapdoors, and the computer controlled the blue trapdoors. The participant's (computer's) goal was that the marble should drop into the bin with as many orange (blue) marbles as possible.

At the start of the experiment, participants received verbal instructions (based on an instruction sheet,

Games — Strategies	BI strategy	EFR strategy
Game 1	C: $a;e$ P: $c;g$	C: $a;e$ P: $d;g$
Game 2	C: $a;e$ P: $c;g$	C: $a;e$ P: $c;g$
Game 3	C: $a;e,b;e,a;f,b;f$ P: $c;g,d;g,c;h,d;h$	C: $a;e,a;f,b;f$ P: $d;g,d;h$
Game 4	C: $a;e,b;e,a;f,b;f$ P: $c;g,d;g,c;h,d;h$	C: $a;e,b;e,a;f,b;f$ P: $c;g,d;g,c;h,d;h$
Game 1$'$	C: e P: $c;g$	C: e P: $c;g$
Game 3$'$	C: e,f P: $c;g,d;g,c;h,d;h$	C: e,f P: $c;g,d;g,c;h,d;h$

Table 1: BI and EFR (FI) strategies for the 6 experimental games in Figures 1 and 2

see Appendix A) regarding the experiment. These instructions emphasized two facts in particular. Participants were advised to regard each game as if it were played against a new opponent. Furthermore, participants were notified about the fact that at the start of each new game, the computer had its strategy planned out, and that strategy was a best response to a strategy that the computer assumed the participant would play.

The participants first played 14 practice games of increasing difficulty so as to get acquainted with the game setting. The experimental phase consisted of 8 rounds. In each round, the participants played the six games that were described above against a computer opponent. The order in which these 6 games were played in each round was randomized. In each new round, the graphical representation of each game was altered, so as to minimize the possibility for participants to recognize the games they played in earlier rounds.

The participants were randomly divided into two groups: Group A and Group B, each consisting of 25 persons. While playing in certain rounds of a game, as they were about to make their choices at their first decision node, Group A participants were asked a multiple-choice question as follows. (i) For games 1-4: "The computer just chose to go [direction computer just chose]. If you choose to go [direction corresponding to playing d], what do you think the computer would do next?", or, (ii) For games $1', 3'$: "It's your turn. If you choose to go [direction corresponding to playing d], what do you think the computer would do next?". Three options were given regarding the likely choice of the computer: "I think the computer would most likely open the left side" or " I think the computer would most likely open the right side" or "Both answers seem equally likely". The first two answers translated to the moves e or f of the computer, respectively. In case of the third answer, we assumed that the participant was undecided regarding the computer's next choice.

For the Group B participants, questions were asked in certain rounds at the end of a game: (i) For games 1-4: "The computer first chose to go [direction computer chose at its first decision point]. When you made your first choice, what did you think the computer would do next if you chose to go [direction corresponding to playing d]?", or, (ii) For games $1', 3'$: "When you made your first choice, what did you think the computer would do next if you chose to go [direction corresponding to playing d]?". The answer choices were the same as above.
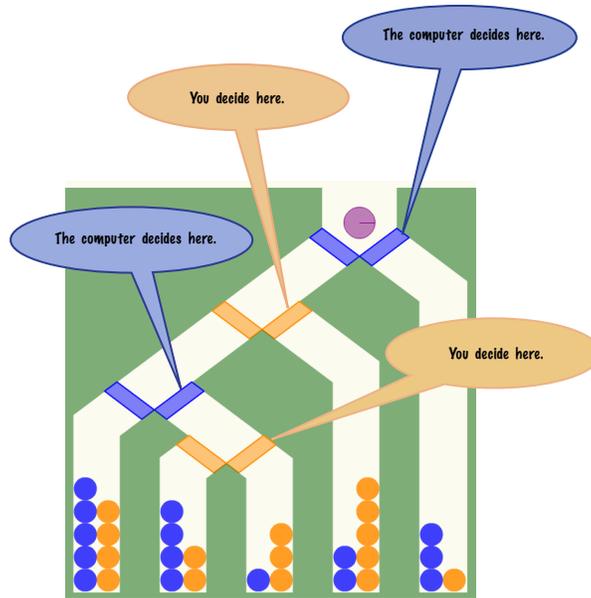
Figure 3: Graphical interface for the participants.

At the end of the experiment, the participants were presented with two questionnaires, both based on an image of one of the experimental games (cf. Figure 4, based on a representation of Game 4). All participants were asked these questions based on the same images. Finally, they were paid according to the marbles they earned at one of the experimental games, selected randomly for each participant by a mechanism proposed by Allais [1], and supported as incentive-compatible by Azrieli et al. [2]. The reward varied from 3.75 euros to 15 euros depending on the number of marbles (1–4) that the participant won in the randomly selected game.

3 Results and analysis

We did not find any significant difference in behavior between participants in groups A and B. Instead, we found strong evidence that participants were equally likely to choose to continue (playing d) at the first decision node (Bayes Factor = 0.056). Consequently, we analyze the data of all 50 participants together.

3.1 Aggregate results: Is there a forward induction trend at participants' first node?

Our main hypothesis was that in case the participants played c more in Game 2 than in Game 1, and played c more in Game 4 than in Game 3, then we could safely assume that the participants' behavior showed some inclination towards FI reasoning. To test this hypothesis, we first perform a mixed effects logistic regression on participant choices in Game 1 and Game 2. In this model, we determine whether the probability that a participant chooses c in a given game depends on whether the game is an instance of Game 1 or an instance of Game 2. To account for individual differences, participants are treated as random effects. The results show that participants were indeed significantly more likely to play c in
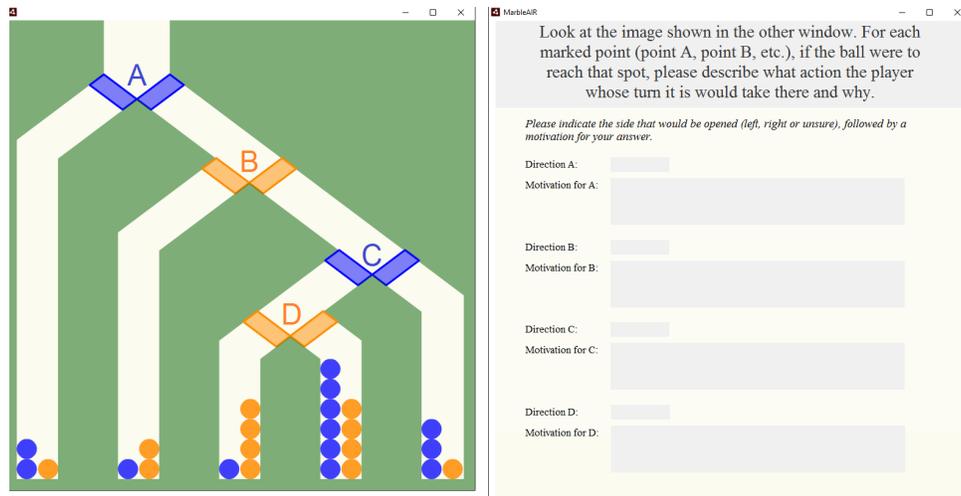
Figure 4: Questionnaire based on a representation of Game 4.

Game 2 than in Game 1 ($p < 0.01$). A second logistic regression for Game 3 and Game 4 shows that that participants were also significantly more likely to play c in Game 4 than in Game 3 ($p < 0.02$). These results suggest that in the aggregate, participants' behavior shows inclination towards FI reasoning.

However, on a visual inspection of the graphs showing the choices of the participants at their first decision nodes (cf. Figure 5) we found the following: Comparing Game 1 to Game 2, only 14 participants out of 50 played c in Game 2 somewhat more than they played c in Game 1 (meaning that the number of times such a participant played c in the rounds of Game 2 minus those he played c in the rounds of Game 1 is more than 1). In contrast, 5 participants played c in Game 2 somewhat less than they played c in Game 1. For the remaining 31 participants, there was not much difference in their playing c between Games 1 and 2. Comparing Game 3 to Game 4, only 10 participants out of 50 played c somewhat more in Game 4 than they did in Game 3 (with no reverse cases), and for the remaining 40 participants there was not much difference between their playing c across Games 3 and 4. We finally note that only 13 participants out of 50 played c in more than half of the 8 rounds of Game $1'$. Interestingly, all these trends in participants' choices at their first decision points indicated a very low extent of game-theoretic strategic reasoning overall—be it BI or FI reasoning.

We should note here that in the current experiment, in all six games, the computer would gain 6 points if the game ended up at the rightmost leaf, compared to only 4 points at the rightmost leaves in the games of [19]. This higher pay-off appears to have motivated the participants to think that the computer would risk for 6 and would continue in the game.

Given these conflicting findings of the logistic regression and the visual inspection of the choice graphs, it is pertinent to ask: "How exactly do participants reason while playing these games?" In what follows, we attempt to answer this question.

3.2 Dividing participants into groups: decisions and answers about opponent

One way to divide the participants into groups is to analyze their answers to the questionnaire presented in Figure 4. According to their answers, we classify the participants along the following three dimensions (see Figure 6): (i) *risk-taking/risk-averse* considerations for self and opponents, (ii) explicit order of
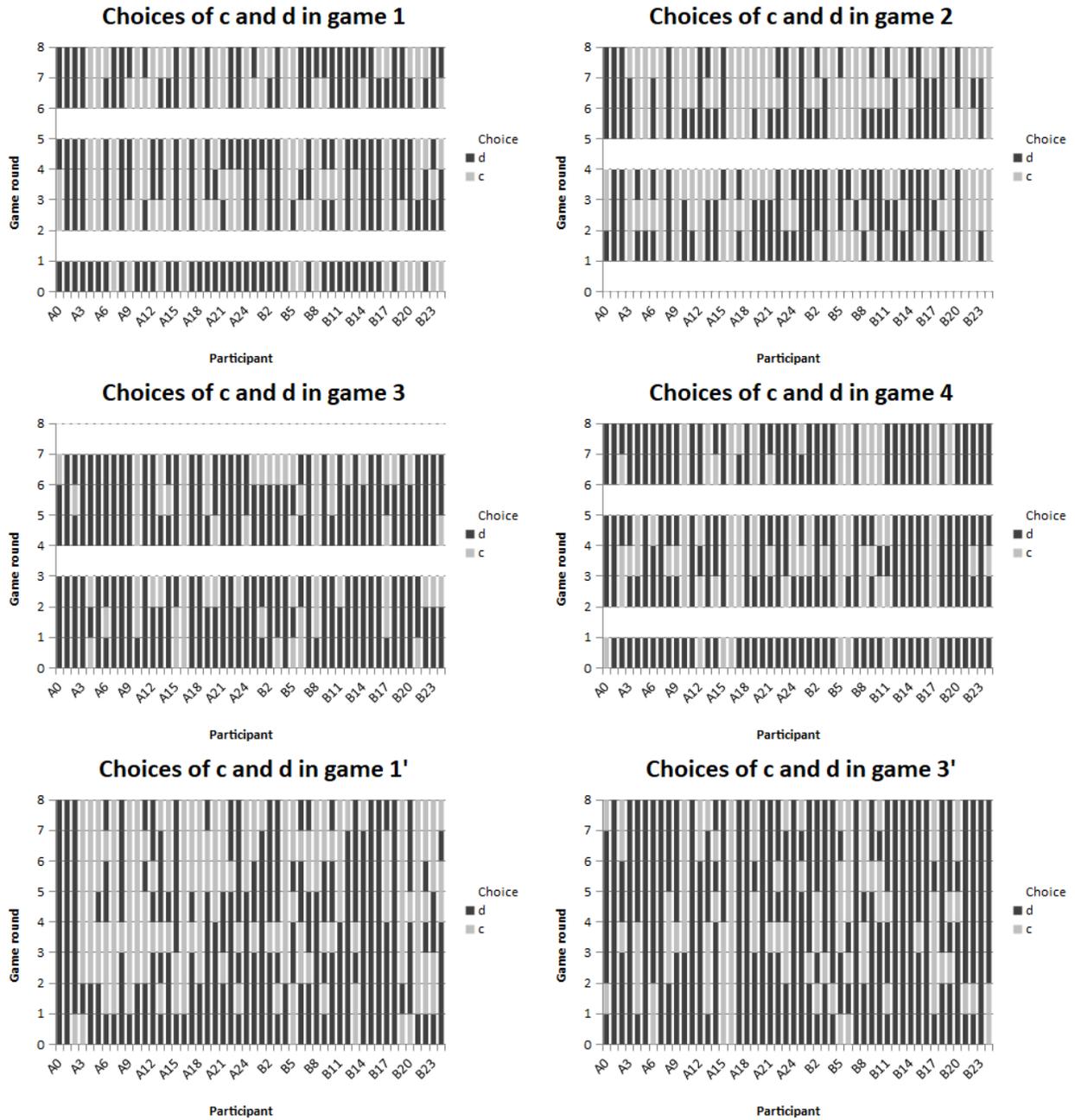
Figure 5: Sequence of choices (across the 8 repetitions of each game) at the first decision node of games 1, 2, 3, 4, $1'$, $3'$, per participant (named A1 ... A25, B1 ... B25). The *dark grey* color corresponds to the rounds the participant played the move *d*, and the *light grey* color corresponds to the rounds the participant played move *c*, whenever the participant's first decision node was reached. Note that white horizontal bands correspond to rounds in which the computer took option *a*, thereby ending the game.

Behavior	Risk-taking (self)	Risk-taking (opponent)	Risk-averse (self)	Risk-averse (opponent)
Participants (Group A)	A0, A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A14, A15, A18, A19, A20, A21, A22, A23, A24	A1, A2, A3, A4, A5, A6, A7, A8, A9, A11, A12, A14, A15, A17, A18, A19, A20, A21, A22, A23, A24	A10, A13, A16, A17	A0, A4, A8, A9, A10, A13, A16, A17, A22
Participants (Group B)	B0, B1, B2, B3, B7, B9, B10, B12, B13, B15, B16, B18, B19, B20, B21, B22, B23, B24	B0, B2, B3, B7, B9, B10, B11, B12, B13, B15, B16, B18, B19, B20, B21, B22, B23, B24	B4, B5, B6, B8, B11, B17	B0, B1, B4, B5, B6, B8, B10, B11, B12, B14, B17, B18, B20, B21

Behavior	Zeroth order theory of mind	First order theory of mind	Second order theory of mind
Participants (Group A)	A0, A1, A3, A5, A8, A13, A15, A19, A20, A22	A2, A4, A7, A9, A10, A12, A14, A16, A17, A21, A23	A6, A11, A18, A24
Participants (Group B)	B3, B11, B14, B15, B16, B20	B0, B2, B4, B5, B6, B7, B8, B9, B10, B12, B13, B17, B18, B19, B21, B22, B24	B1, B23

Behavior	Competitive	Cooperative	Neither
Participants (Group A)	A0, A1, A2, A3, A10, A14, A15, A16, A19, A20, A21	A5, A7, A11, A17, A23, A24	A4, A6, A8, A9, A12, A13, A18, A22
Participants (Group B)	B3, B5, B14, B22	B0, B1, B2, B7, B8, B9, B11, B13, B16, B18, B19, B21, B23, B24	B4, B6, B10, B12, B15, B17, B20

Figure 6: Reasoning behind participant behaviour. Answers to final question classified on three dimensions: attitude towards risk (left); explicit order of theory of mind (middle) and level of cooperation (right)

theory of mind, which has been shown to be important in strategic reasoning in dynamic games [32, 33], and (iii) *competitive/cooperative* considerations. To come up with our classification, we considered the decisions at each node of the game in Figure 4 in the questionnaire and the participants' explanations of reasoning behind their decisions.

As an example, consider the following sample answers of the participants A7 and B0 to the questionnaire given in Figure 4:[1]

Participant A7:
Direction A: right
Motivation: B would probably open right, at which point C has the chance to get at least 3 marbles.
Direction B: right
Motivation: If C opens left, B has a 100% chance of getting 4 marbles. The chance that C opens left is big enough, since there are 6 marbles in D-right.
Direction C: left
Motivation: Since the direction for D doesn't matter, and the amount of marbles of your opponent doesn't matter, the chance is quite big that D will choose right, giving you 6 marbles.

---

[1] The full answers of all participants' answers to the final questionnaire will be provided on `http://www.ai.rug.nl/SocialCognition/experiments/`.
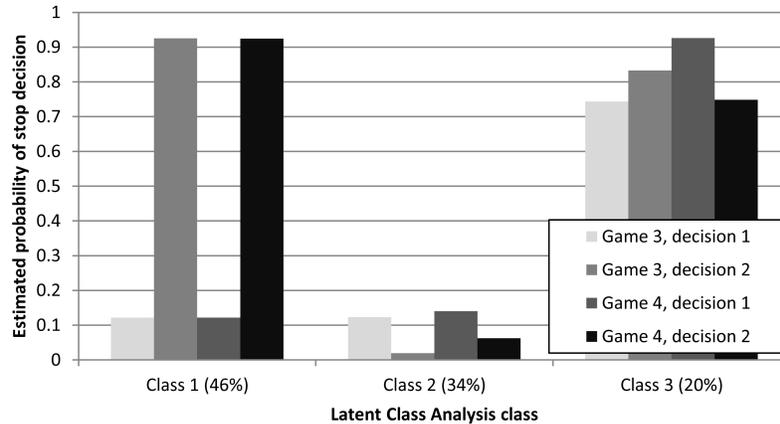
Figure 7: Estimated probability of stop decision (i.e. *c* or *g*) according to the latent class analysis of participant choices on Game 3 and Game 4.

Direction D: right
Motivation: I would pick right. Even against the computer it felt wrong to pick left (since the opponent's score doesn't matter for my own score).

Participant B0:
Direction A: right
Motivation: The player would hope for the largest amount of points he could get. 2 would not be enough when there is the chance to get three or six.
Direction B: right
Motivation: The player would chose right because there his/her chances are good that he/she will end up with 4 balls. But there is the risk of having only one ball
Direction C: right
Motivation: If the player decides to go left, he will let the other person choose how many balls he will end up with. This is a high risk since it is possible to get just one ball. Therefore the player would take the right path were he has the mean of 3.
Direction D: right
Motivation: The player would chose right since both boxes contain 4 balls. Further it is not important to let the other player have less balls than you. Therefore I would say the player is nice and let the other person have the full amount of balls.

Based on these answers, we came up with the classification shown in Figure 6. Participant A7 considers the chance of getting more for both the players (risk-taker: self and opponent), takes into consideration how the opponent would play (first-order theory of mind), and does not want to wrong the opponent (cooperative). Participant B0 also considers the chance of getting more for both the players (risk-taker: self and opponent), in addition, considers safe play for the opponent (risk-averse: opponent), takes into consideration how the opponent would play (first-order theory of mind), and wants to be nice to the opponent (cooperative). Note that Directions A and C in Figure 4 (blue trapdoors) are considered to be the computer's moves, and Directions B and D (orange trapdoors) are considered to be the participant's moves in the interpretation here.

3.3   Dividing the participants by latent class analysis according to their decisions

To further analyze participant choices, we used latent class analysis (LCA, [29]), allowing us to divide participants into classes that exhibit similar patterns of behavior. In our model, the behavior of a participant is described by whether they choose c or d at their first decision point of a game, and whether they choose g or h at their second decision point. LCA determines what classes of behavior best describe the observed participant data, given the number n of such classes. Statistical model selection tools can be used to select the number of classes that is most appropriate for the data.

Our aim of the latent class analysis is to describe the behavior of participants in both their decision points. However, of the 181 times (out of 300 times where the participants actually reached their first decision node) that a participant in Game 1 ended up at his or her second decision point, only 2 times participants chose h. In addition, no participant in Game 2 ever reached the second decision point—the computer was programmed to play e whenever its second decision node was reached. For this reason, we performed a latent class analysis on the two decision points of Game 3 and Game 4 only. The Bayesian Information Criterion (BIC) favors an LCA model with three classes, which we describe here. The results of the estimation parameters of the LCA are summarized in Figure 7. In this figure, estimated probabilities of choosing c at the first decision point or g at the second decision point are averaged over repeated encounters of the same game. According to the LCA, 46% of the participants tend to choose to continue at the first decision point, but stop at the second decision point: Class 1. A further 34% of the participants is represented by the second class, who prefer to continue at both decision points: Class 2. The final 20% of the participants has a preference to stop at both decision points: Class 3.

Thus, about 80% of the participants belonged to Class 1 or 2, who tend to continue at the first decision node, and the remaining 20% tend to stop. Out of these 41 participants assigned to Class 1 or Class 2 by LCA, 37 were classified as risk-takers (self), 37 were classified as risk-takers (opponent), while only two participants were classified as neither type of risk-taking based on their final answers (see Figure 6). Out of the 9 participants assigned to Class 3 by LCA, 7 were classified as risk-averse (self), 4 were classified as risk-averse (opponent), while only one participant did not mention any risk-aversion, once again, based on their final answers.

As depicted in Figure 1, the final decision point for Player P in Game 3 and Game 4 only influences the payoff for the computer player C. Participants that ended up at this decision point more often chose the competitive option (60%), in which the opponent only received a payoff of 1, than the cooperative option (40%), which yielded the opponent a payoff of 6 (Bayes Factor = 132.7663). In addition, of the 19 participants assigned to Class 2 by the LCA, who are therefore classified as being likely to choose the cooperative option at their second decision point, 15 were also classified as being cooperative based on their final answer (see Figure 6). Moreover, of the 31 remaining participants, only 5 were classified as being cooperative based on their final answer. However, there was no clear mapping of classes 1 and 3 of the LCA results and the classification 'competitive' and 'neither cooperative nor competitive' based on the final answers.

3.4   Individual participant decision patterns: differences and similarities

Visual inspection of individual participant graphs shows a similar pattern. See Figure 8 for an example of a graph of an individual participant's choices: In two of the rounds for games 1-4, selected randomly, the computer took the *outside option*, that is, played a, and so the first decision node of the participant was not reached. It was noted that participants do differentiate between Game 1 and Game 2 on the one hand and Game 3 and Game 4 on the other hand: The participants played c more often in Games
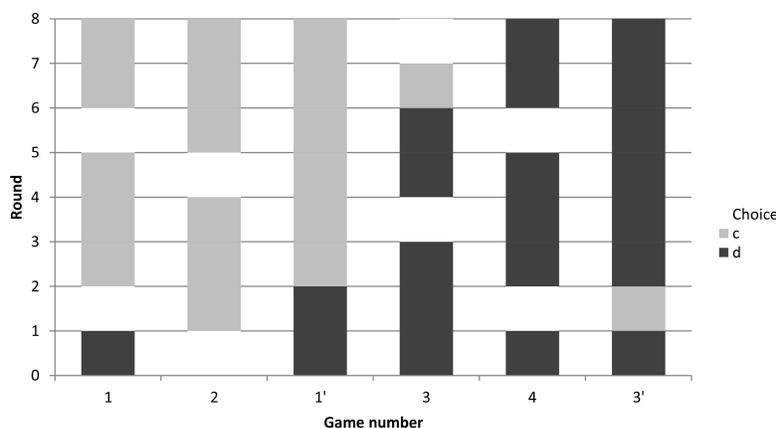
Figure 8: Example of participant choices at the first decision point across all games. This participant, like many others, tended to choose *c* at the first decision point for Game 1, Game 2, and Game 1′, but tended to choose *d* for Game 3, Game 4, and Game 4′.

1 and 2, compared to what they played in Games 3 and 4. Further, even though participants' behavior differed between Game 1 and Game 2 to some extent (not exceedingly so), participants appeared to behave similarly in Game 3 and Game 4. For the truncated games, behavior in Game 1′ appeared similar to Game 1, and behavior in Game 3' appeared similar to both Game 3 and Game 4.

4 Discussion

The results outlined in Section 3 indicate that while in the aggregate, participant behavior is indicative of forward reasoning, most participants do not appear to making use of either forward induction reasoning or backward induction reasoning. Apparently, game-theoretic rationality does not drive participants' choices in these turn-taking games. In this section, we take a look at some possible alternatives for the driving force behind people's choices in our experiment, in the light of the literature on human decision making and perspective taking in games.

4.1 Level-*k* reasoning

In our experiment, we explicitly mention to the participants that the computer player acts based on its beliefs about future participant behavior, and that the computer player does not learn from the participant's previous actions. Participants are therefore playing a series of one-shot games against a computer opponent. In behavioral economics, the behavior of participants in such one-shot games has been successfully modeled through iterated best-response models such as level-*n* theory [3, 4, 14, 15, 36, 47], cognitive hierarchies [9], quantal response equilibria [31], and noisy introspection models [23].

Similar to the theory of mind classification we describe in Section 3, a participant's level of reasoning sophistication is measured by the maximum number of steps of iterated reasoning the participant considers. According to cognitive hierarchy models [9], a naive level-0 reasoner does not reason strategically at all, but instead chooses randomly among all available options. In contrast, a level-1 reasoner believes that all other participants are level-0 reasoners, and selects the option that is a best response to that belief. In terms of our Marble Drop game, level-1 reasoners would therefore maximize future payoffs without

considering the opponent's past behavior. In the cognitive hierarchy model, a level-2 reasoner performs two steps of iterated reasoning. Such a reasoner believes that other individuals can be either level-0 reasoners that play randomly, or level-1 reasoners that maximize their expected payoff. In addition, level-2 reasoners form beliefs about the relative proportion of level-0 reasoners and level-1 reasoners.

Over a range of one-shot non-repeated games, cognitive hierarchy models estimate participants to be level 1.5 reasoners on average [9, 13]. In terms of theory of mind reasoning, the vast majority of participants in these one-shot non-repeated games reason at zero-order or first-order theory of mind. Although part of the participants in those experiments were found to use more than two steps of iterated reasoning, only few players were found to be well-described as higher-level agents [51].

4.2   Considering the opponent's perspective

Kawagoe has recently applied a level k-analysis to centipede games to explain that participants' lack of backward induction may be due to their believing that the opponent could have made an error or cannot apply backward induction for the number of steps required [28]. Evans and Krueger have shown in a very simple dynamic trust game of perfect information, in which participants need to open windows in order to inspect payoffs, quite a few participants do not choose to inspect all the opponent's payoffs but they do inspect all their own ones, thereby not applying even first-order theory of mind [16, 17]. For higher orders of theory of mind, the situation seems even worse. It has also been shown that level k reasoning for higher k is correlated to general cognitive ability [22].

These results could make one quite pessimistic about the value and actual use of higher-level perspective-taking in games. Is the situation really that bad? De Weerd and colleagues have shown on the basis of agent simulations that both in competitive situations and in repeated mixed-motive interactions, agents attain higher pay-offs if they apply second-order theory of mind than first-order theory of mind, which is in turn more beneficial than zero–order theory of mind [48, 50], so the application of theory of mind is beneficial. Also theoretically, higher-order theory of mind is required in dynamic perfect information games [38]. But do people really use it? De Weerd and colleagues [48] showed that, by letting people play unknowingly against a second-order theory of mind agent in a repeated negotiation game, people are enticed to apply second-order theory of mind, which became much more prevalent than when the participants played against zero-order or first-order agents.

In cognitive science, the application of perspective taking in turn-taking games of perfect information has been studied, especially on the basis of experiments in which the decision trees looked like our Game $1'$ and $3'$, but with a large variation of pay-off structures, not only those similar to centipede games. These various turn-taking games, when solved by participants without knowledge of the backward induction algorithm, do require second-order theory of mind: "the opponent *thinks* that at my final move, I *intend* to go down". Hedden and Zhang [25] showed that people have a hard time learning to apply second-order in these games (around 60 % optimal decisions after many game items). Meijering and colleagues [32, 33] proposed several supportive interventions that helped people to make much better decisions (up to around 90 % optimal ones at the end of playing many experimental games with different payoff structures).

Even though application of second-order theory of mind can apparently be supported and trained, Meijering and colleagues [35] also showed on the basis of computational cognitive models that people do not start to apply higher orders of theory of mind spontaneously. Rather, people start to do this only when they receive negative feedback from the results of the games they play, for example, in the form of low pay-off, or the comment "you could have chosen better". They prefer reasoning that is "as easy as possible, as complex as necessary". And even when they do produce the backward induction *outcome*, it appears by analyzing their eye movements and reaction times that their reasoning rather follows a stream

of thinking starting from the root to its child nodes and then to the grandchildren in the decision tree, by "forward reasoning plus backtracking", and giving special attention to decision points [7, 34].

Coming back to the experiment

The experiments done by [32–34] concern games in which the computer opponent has been programmed to make rational decisions, with the goal to attain the highest possible individual pay-off, and participants are told about this. In the games that participants play in the current paper, in contrast, the opponent often plays irrationally, and participants are told that in fact, the computer opponent is playing its best response to a strategy that it attributes to the participant. Let us see whether this leads to a difference in the application of theory of mind. For our previous experiment presented in [19], we found that higher levels of theory of mind expressed in final answers was correlated positively to number of marbles won in total in the games [21], so it appears that theory of mind is indeed profitable in this type of Marble Drop games in which the opponent sometimes acts in an unexpected way by not choosing to end the game.

In their model-based analysis of strategies used by the participants of [18, 33], Meijering and colleagues [35] implemented one simple "zero-order theory of mind" strategy that ignores any future decisions and simply compares the immediate payoff, when stopping a game, against the maximum of all future possible payoffs, which they dubbed as a "simple risk-taking strategy". A more advanced "first-order theory of mind" version of that model attributes the simple risk-taking strategy to the opponent. They showed that the sequences of decisions of many 9-year old children playing the games of [18] correspond to either the zero -order risk-taking strategy or the first-order version attributing risk-taking to the opponent; and that the adults' decisions in these games [33] correspond more closely to the first-order strategy of assigning the risk-taking strategy to the opponent, or to a second-order strategy.

It is interesting to see in Figure 6 that in our games, many participants' answers also exemplify that they themselves "take a risk" or "take a chance" when moving to the right, namely 40 out of 50 participants; and a similar number of participants, namely 39 out of 50, attributes such risk-taking to their opponent. Also, in line with the work by Meijering and colleagues, most of our current adult participants (34 out of 50) appear to reason explicitly at the first or second order of theory of mind, at least when their answers are analyzed. Thus, it does appear that participants regularly do think and act at least at first- or second-order theory of mind, which is rather high when considering how much more complicated our decision trees for games 1, 2, 3 and 4 are than those of [25, 33] and especially [16].

For the current experiment, it would be useful to estimate more precisely which level of theory of mind each participant actually applies when making their own decisions, using computational techniques such as simulation of computational cognitive models [20, 21] or Bayesian strategy estimation [49].

4.3 Moving toward cooperation

It has long been noted in the literature on behavioral game theory that players' utilities do not correspond one-to-one to payoffs [30]. Healy [24] has elicited players' utilities over outcomes associated with centipede-like games, showing that there are many motivations that play a role. In our games, trying to get all the way to the right also corresponds to going for the outcomes with maximum social welfare, see Figures 1 and 2. It is very well possible that participants try to get there and aim to entice their opponent to move right as well (see also Nagel and Tang's experiment [37], in which participants may have reason to believe that their opponent could be an altruist). In this light, our participants could

interpret the opponent's first "irrational" move to the right not as "irrational" at all, but as a first step towards cooperation to attain maximum social welfare.

5   Conclusion

We made a number of improvements that we thought would make it easier for participants to apply backward or forward induction reasoning in the perfect-information games in a Marble Drop set-up than in [19]. We found that even though in the aggregate, participants in the new experiment still tend to slightly favor the forward induction choice at their first decision node, their verbalized strategies most often depend on their own attitudes towards risk and those they assign to the computer opponent, sometimes in addition to considerations about cooperativeness and competitiveness.

When we analyzed the individual participants' decisions and their answers about the reasons behind their decisions, it turns out that many players do not find anything to rationalize when the computer does not take the safe and rational option to go down and stop the game at its first decision point. Instead, in almost all the games, many participants think that the computer is just taking a chance to gain a higher payoff later. Also, many of the players are willing to take such a risk at their own first decision point: very few players take the BI option of stopping the game there. Dividing the players into classes according to their decisions as well as according to how they reasoned about their own and the opponent's choices turned up several more nuanced patterns of reasoning.

In view of our findings for this experiment, our subsequent categorization of the elements of reasoning behind the behavior of participants, namely (i) risk-taking/risk-averseness tendency of self/opponent, (ii) competitive/cooperative tendency, and (iii) proficiency in applying higher-order theory of mind, has provided an understanding of a few possible player-types with respect to their strategic reasoning. In addition, the distinction of instinctive versus contemplative reasoners could be addressed in future investigations, requiring a detailed analysis of the time data [44, 45]. Such an analysis of the temporal data could also provide new insights into the categorisation of the participants' behavior as mentioned above.

The take-home message from this paper that we would like to suggest is the following: In addition to investigating into game-theoretic rationality as a guiding force for understanding people's choices in dynamic games (both perfect and imperfect information games), one should also look into risk-taking and risk-averse behaviours, theory of mind considerations, competitive and cooperative considerations, instinctive and contemplative behaviours, and similar reasoning tendencies towards providing a better explanation of people's choices in turn-taking games.

References

[1] M. Allais (1953): *Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine*. Econometrica 21, pp. 503–546.

[2] Y. Azrieli, C. P. Chambers & P. J. Healy (2012): *Incentives in experiments: A theoretical analysis*. Working paper.

[3] M. Bacharach & D. O. Stahl (2000): *Variable-frame level-n theory*. Games and Economic Behavior 32(2), pp. 220–246, doi:10.1006/game.2000.0796.

[4] D. Balkenborg & R. Nagel (2016): *An Experiment on Forward vs. Backward Induction: How Fairness and Level k Reasoning Matter*. German Economic Review 17(3), pp. 378–408, doi:10.1111/geer.12099.

[5] P. Battigalli (1996): *Strategic rationality orderings and the best rationalizability principle*. Games and Economic Behavior 13, pp. 178–200, doi:10.1006/game.1996.0033.

[6] P. Battigalli (1997): *On rationalizability in extensive games*. Journal of Economic Theory 74, pp. 40–61, doi:10.1006/jeth.1996.2252.

[7] G. Bergwerff, B. Meijering, J. Szymanik, R. Verbrugge & S. Wierda (2014): *Computational and algorithmic models of strategies in turn-based games*. In: Proceedings of the 36th Annual Conference of the Cognitive Science Society, pp. 1778–1783.

[8] G. P. Cachon & C. Camerer (1996): *Loss-avoidance and forward induction in experimental coordination games*. The Quarterly Journal of Economics 111(1), pp. 165–94, doi:10.2307/2946661.

[9] C. F. Camerer, T.-H. Ho & J.-K. Chong (2004): *A cognitive hierarchy model of games*. The Quarterly Journal of Economics 119(3), pp. 861–898, doi:10.1162/0033553041502225.

[10] J. Chen & S. Micali (2011): *The robustness of extensive-form rationalizability*. Working paper.

[11] J. Chen & S. Micali (2013): *The order independence of iterated dominance in extensive games*. Theoretical Economics 8, pp. 125–163, doi:10.3982/TE942.

[12] N. Chlaß & A. Perea (2016): *How do people reason in dynamic games?* Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Economic Theory and Personality, No. A04-V2.

[13] M. A. Costa-Gomes & V. P. Crawford (2006): *Cognition and behavior in two-person guessing games: An experimental study*. The American Economic Review 96(5), pp. 1737–1768, doi:10.1257/aer.96.5.1737.

[14] M. A. Costa-Gomes, V. P. Crawford & B. Broseta (2001): *Cognition and behavior in normal-form games: An experimental study*. Econometrica 69(5), pp. 1193–1235, doi:10.1111/1468-0262.00239.

[15] V. P. Crawford, M. A. Costa-Gomes & N. Iriberri (2013): *Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications*. Journal of Economic Literature 51(1), pp. 5–62, doi:10.1257/jel.51.1.5.

[16] A. M. Evans & J. I. Krueger (2011): *Elements of trust: Risk and perspective-taking*. Journal of Experimental Social Psychology 47(1), pp. 171–177, doi:10.1016/j.jesp.2010.08.007.

[17] A. M. Evans & J. I. Krueger (2014): *Outcomes and expectations in dilemmas of trust*. Judgment and Decision making 9(2), p. 90.

[18] L. Flobbe, R. Verbrugge, P. Hendriks & I. Krämer (2008, to appear): *Children's Application of Theory of Mind in Reasoning and Language*. Journal of Logic, Language and Information. Special issue on formal models for real people, edited by M. Counihan.

[19] S. Ghosh, A. Heifetz & R. Verbrugge (2015): *Do players reason by forward induction in dynamic perfect information games?* In R. Ramanujam, editor: Proceedings Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2015, Carnegie Mellon University, Pittsburgh, USA, June 4-6, 2015., EPTCS 215, pp. 159–175, doi:10.4204/EPTCS.215.12. Available at https://doi.org/10.4204/EPTCS.215.12.

[20] S. Ghosh, B. Meijering & R. Verbrugge (2014): *Strategic Reasoning: Building Cognitive Models from Logical Formulas*. Journal of Logic, Language and Information 23(1), pp. 1–29.

[21] S. Ghosh & R. Verbrugge (2017): *Studying strategies and types of players: Experiments, logics and cognitive models*. Synthese, pp. 1–43, doi:10.1007/s11229-017-1338-7.

[22] D. Gill & V. Prowse (2016): *Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis*. Journal of Political Economy 124(6), pp. 1619–1676, doi:10.1086/688849.

[23] J. K. Goeree & Ch. A. Holt (2004): *A model of noisy introspection*. Games and Economic Behavior 46(2), pp. 365–382, doi:10.1016/S0899-8256(03)00145-3.

[24] P. J. Healy (2016): *Epistemic experiments: Utilities, beliefs, and irrational play*. Available at http://healy.econ.ohio-state.edu/present/Healy-EpistemicPresentation90min.pdf. Unpublished presentation.

[25] T. Hedden & J. Zhang (2002): *What do you think I think you think?: Strategic reasoning in matrix games*. Cognition 85(1), pp. 1–36, doi:10.1016/S0010-0277(02)00054-9.

[26] A. Heifetz & A. Perea (2015): *On the outcome equivalence of backward induction and extensive form rationalizability*. International Journal of Game Theory 44(1), pp. 37–59, doi:10.1007/s00182-014-0418-x.

[27] S. Huck & W. Müller (2005): *Burning money and (pseudo) first-mover advantages: An experimental study on forward induction*. Games and Economic Behavior 51(1), pp. 109–127, doi:10.1016/j.geb.2004.03.006.

[28] T. Kawagoe & H. Takizawa (2012): *Level-k analysis of experimental centipede games*. Journal of Economic Behavior and Organization 82(2), pp. 548–566, doi:10.1016/j.jebo.2012.03.010.

[29] A. L. McCutcheon (1987): *Latent Class Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences 07-064, Sage Publications, Newbury Park, CA.

[30] R. D. McKelvey & T. R. Palfrey (1992): *An experimental study of the centipede game*. Econometrica: Journal of the Econometric Society, pp. 803–836, doi:10.2307/2951567.

[31] R. D. McKelvey & T. R. Palfrey (1995): *Quantal response equilibria for normal form games*. Games and Economic Behavior 10(1), pp. 6–38, doi:10.1006/game.1995.1023.

[32] B. Meijering, L. van Maanen, H. van Rijn & R. Verbrugge (2010): *The facilitative effect of context on second-order social reasoning*. In: Proceedings of the 32nd Annual Conference of the Cognitive Science Society, pp. 1423–1428.

[33] B. Meijering, H. van Rijn, N. A. Taatgen & R. Verbrugge (2011): *I do know what you think I think: Second-order theory of mind in strategic games is not that difficult*. In: Proc. 33rd Annual Conf. Cognitive Science Society, pp. 2486–2491.

[34] B. Meijering, H. van Rijn, N.A. Taatgen & R. Verbrugge (2012): *What eye movements can tell about theory of mind in a strategic game*. PloS ONE 7(9), p. e45961, doi:10.1371/journal.pone.0045961.

[35] B. Meijering, N. A. Taatgen, H. van Rijn & R. Verbrugge (2014): *Modeling inference of mental states: As simple as possible, as complex as necessary*. Interaction Studies 15(3), pp. 455–477, doi:10.1075/is.15.3.05mei.

[36] R. Nagel (1995): *Unraveling in guessing games: An experimental study*. The American Economic Review 85(5), pp. 1313–1326.

[37] R. Nagel & F. F. Tang (1998): *Experimental results on the centipede game in normal form: An investigation on learning*. Journal of Mathematical Psychology 42(2), pp. 356–384.

[38] E. Pacuit (2015): *Dynamic models of rational deliberation in games*. In J. van Benthem, S. Ghosh & R. Verbrugge, editors: Models of Strategic Reasoning: Logics, Games and Communities, *LNCS-FoLLI* 8972, Springer, New York, NY, pp. 3–33.

[39] D. Pearce (1984): *Rationalizable strategic behaviour and the problem of perfection*. Econometrica 52, pp. 1029–1050, doi:10.2307/1911197.

[40] A. Perea (2007): *Epistemic foundations for backward induction: An overview*. In J. van Benthem, D. Gabbay & B. Löwe, editors: *Interactive Logic*, Texts in Logic and Games 1, Amsterdam University Press, pp. 159–193.

[41] A. Perea (2012): *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press, Cambridge, doi:10.1017/CBO9780511844072.

[42] D. Premack & G. Woodruff (1978): *Does the chimpanzee have a theory of mind?* Behavioral and Brain Sciences 1(04), pp. 515–526, doi:10.1017/S0140525X00076512.

[43] P. J. Reny (1992): *Backward induction, normal form perfection and explicable equilibria*. Econometrica 60, pp. 627–649, doi:10.2307/2951586.

[44] A. Rubinstein (2013): *Response time and decision making: An experimental study*. Judgment and Decision Making 8(5), pp. 540–551, doi:10.1.1.396.668.

[45] A. Rubinstein (2016): *A typology of players: Between instinctive and contemplative*. Quarterly Journal of Economics 131(2), pp. 859–890, doi:10.1093/qje/qjw008.

[46] Q. Shahriar (2014): *An experimental test of the robustness and the power of forward induction*. Managerial and Decision Economics 35, pp. 264–277, doi:10.1002/mde.2613.

[47] D. O. Stahl & P. W. Wilson (1995): *On players' models of other players: Theory and experimental evidence*. Games and Economic Behavior 10(1), pp. 218–254, doi:10.1006/game.1995.1031.

[48] H. de Weerd, R. Verbrugge & B. Verheij (2017): *Negotiating with other minds: The role of recursive theory of mind in negotiation with incomplete information*. Autonomous Agents and Multi-Agent Systems 31(2), pp. 250–287, doi:10.1007/s10458-015-9317-1.

[49] H. de Weerd, D. Diepgrond & R. Verbrugge (2017, to appear): *Estimating the use of higher-order theory of mind using computational agents*. The B.E. Journal of Theoretical Economics.

[50] H. de Weerd, R. Verbrugge & B. Verheij (2013): *How much does it help to know what she knows you know? An agent-based simulation study*. Artificial Intelligence 199, pp. 67–92, doi:10.1016/j.artint.2013.05.004.

[51] J. R. Wright & K. Leyton-Brown (2010): *Beyond equilibrium: Predicting human behavior in normal-form games*. In: Proceedings of the 24th Conference on Artificial Intelligence, pp. 901–907.

Appendix A: Instruction sheet for the participants

Here follows the written information on the instruction sheet that was given to each participant and explained by the experimenter before the start of the training games.

- In this task, you will be playing two-player games. The computer is the other player.

- In each game, a purple marble is about to drop, and both you and the computer determine its path by controlling the orange and the blue trapdoors.

- You control the orange trapdoors, and the computer controls the blue trapdoors.

- Your goal is that the purple marble drops into the bin with as many orange marbles as possible. The computer's goal is that the marble drops into the bin with as many blue marbles as possible.

- Click on the left trapdoor if you want the marble to go left, and on the right trapdoor if you want the marble to go right.

- How the computer reasons in each particular game?

  The computer thinks that you already have a plan for that game, and it plays the best response to the plan it thinks that you have for that game.
  However, the computer does not learn from previous games and does not take into account your choices during the previous games.

- The first 14 games are practice games. At the end of each practice game, you will see how many marbles you collected in that game.

- The practice games are followed by 48 experiment games. At the end of each experiment game, you will see how many marbles you collected in that game.

- You will be able to start each game by clicking on the "START GAME" button, and move to the next game by clicking on the "NEXT" button.

- At some points during the experiment phase, you will be asked a few questions regarding your choices. Please think carefully about each question before answering it.

- There will be a break of 5 minutes once you finish 24 of the 48 experiment games.

- After the the 48 experiment games, you will be asked a few final questions regarding 2 example games.

- The money you will earn is between € 3.75 and € 15 and depends on the marbles you have won during the experiment phase. At the end of the experiment, 1 of the 48 experiment games will be chosen at random. Games where you were not given at least one choice will be omitted from this random selection. The number of marbles that you have collected in the randomly selected game will determine your final earnings: each collected marble is worth € 3.75.