ORDER OF THEORY OF MIND IN THE REVERSED BIDDING GAME

Bachelor's Project Thesis

Tim Haarman, s2404184, t.r.haarman@student.rug.nl

Supervisor: dr. H.A. de Weerd

Abstract: People use theory of mind (*ToM*) to reason about the beliefs of others. This reasoning can be done recursively, with the different levels being called the order of *ToM*. First-order *ToM* equates to only reasoning about another's beliefs, whereas second-order *ToM* is reasoning about the beliefs someone has about the beliefs of another, and so forth. In many studies it is argued that people use second-order *ToM* or higher in strategic games. In this paper, people played a strategic *ToM* game called 'reversed bidding' against a computer. The goal for the participants was to make a bid that was as high as possible, while being lower than what they thought the computer was going to bid. The participant choices were compared to the predictions of different orders of *ToM* agents. It was found that the results from the participants resembled the predictions of the zero-order *ToM* agent the closest, and showed less resemblance with higher orders of *ToM*. This suggests that people do not always use *ToM* in strategic situations.

1. Introduction

Theory of mind is the ability of people to create a model of the mental content of others, such as their beliefs, intentions or desires. This can be used to predict what another person will do (Premack & Woodruff, 1978). An example of this is "Bob thinks that Anne is hungry". Here, Bob reasons about the mental state of Anne, and can use this belief to predict that Anne will for example get some food. It is also possible for people to nest these mental models, resulting in higher order theory of mind. An example of such a nested belief is "Bob thinks that Anne thinks that Bob is hungry". In this example, second-order theory of mind is used, as the reasoning is two levels deep. Earlier research has shown that people are able to use this kind of reasoning while playing strategic games (Hedden & Zhang, 2002; Meijering, Van Rijn, Taatgen & Verbrugge, 2011). It has furthermore been argued that people can gain an advantage from using higher order theory of mind in a competitive setting, such as competitive strategic games (De Weerd, Verbrugge & Verheij, 2013). However, it is also argued that people only use theory of mind up until a certain order (Flobbe, Verbrugge, Hendriks, & Krämer, 2008). The aim of this paper is to find out whether and to what extent people in reality use theory of mind when faced with a simple competitive strategic game, in the form of the reversed bidding game. This game is explained in Section 2. To see whether people use theory of mind and if so, what order of theory of mind is most prevalent, the participant behaviour is compared to the predictions of computational models in the form of theory of mind agents. These agents are introduced by De Weerd et al. (2013), and will be discussed in Section 3.2. The participant behaviour is furthermore compared to three simple strategies, which will be explained in Section 3.4. The simulated outcomes of the multiple different orders of these theory of mind agents and the three simple strategies will be laid out next to the participant data to compare and see what order of theory of mind was most likely used by the participants that partook in the experiment.

2. Reversed Bidding game

To determine to what extent people use theory of mind during a simple strategic game, the predictions of computational theory of mind agents are compared to the data obtained from participants in the setting of the reversed bidding game. The reversed bidding game is in essence a very simple adaptation of a standard closed bidding game, where two or more people have to try to be the highest bidder, but of course want to keep the costs as low as possible for themselves. The optimal bid would be to be a little bit above the highest bid of the other player(s).

Table 2.1: The pay-off matrix of the reversed bidding game. The table rows show the pay-offs for the player choosing the row bid, for every possible choice of the opponent choosing the column bid.

	1	2	3	4	5	6	7
1	0	6	5	4	3	2	1
2	0	0	7	6	5	4	3
3	0	0	0	8	7	6	5
4	0	0	0	0	9	8	7
5	0	0	0	0	0	10	9
6	0	0	0	0	0	0	11
7	0	0	0	0	0	0	0

The reversed bidding game changes this. As the name already suggests, it is the same concept, but in reverse. The reversed bidding game is a round-based bidding game that is played by two players, where each round both players simultaneously make a bid. The goal of the reversed bidding game is to score as many points as possible, which can be achieved by correctly guessing what the opponent will do. With a higher bid the maximum score a player could obtain is also higher, but only if the bid that the opponent made was higher than the bid of the participant. If the opponent's bid was either lower or equal to the participant's bid, the participant would get zero points. This resulted in the pay-off matrix as shown in Table 2.1, which was also displayed to the participants during the game.

This game is selected for this experiment, as it provides an interesting setting to study the use of theory of mind in participants. In the reversed bidding game the optimal choice a participant can make is dependent on what their opponent will choose. Because of this, the participants have an incentive to reason about what their opponent will do, which makes it possible for the participants to use theory of mind to their advantage. This specific adaptation in the form of the reversed bidding game is being used, as it is unlikely that participants have had to do this type of bidding before. Earlier experience may influence the way participants make use of theory of mind, which is avoided by choosing for an adapted form such as the reversed bidding game. Analyzing these choices could give new insights into how people reason in a situation where one has to think about the intentions of another player.

3. Methods

3.1. Participants

The game was played by twenty-nine participants over the course of two weeks. It was played by participants of both gender, of which there were nineteen male and ten female participants. Twenty-six of them had studied or were at the time of participation studying at a university. The age of the participants varied from eighteen to fifty-six years (mean: 24.2). All participants gave informed consent for this study prior to participation.

3.2. Theory of mind agents

The agents that were used in this paper to compare the data of the participants to are the theory of mind agents from De Weerd et al. (2013) with slight adaptations to implement them in the reversed bidding game. A basic introduction to these agents will be given, a more technically in depth insight in the workings and programming of these agents can be found in De Weerd et al. (2013).

The zero-order theory of mind agent (ToM_0) is not capable of reasoning about the beliefs of its opponent. It doesn't take into account that the other player may also want to win, and does not reason about the opponent's pay-off matrix. It can however look at its own pay-off matrix, and to the history of the game. Based on that, it will make its decisions. It is only looking to maximize its own score. For example, if an opponent in the reversed bidding game would continuously bid 5, a zero-order theory of mind agent would notice this, and have an increased likelihood of choosing 4 the next round. It forms its beliefs purely on the history of the game.

A first-order theory of mind agent (ToM_1) can go beyond that, as it is also capable of placing itself in the position of the opponent. During a round, it can reason about what it would do if it were in the situation of its
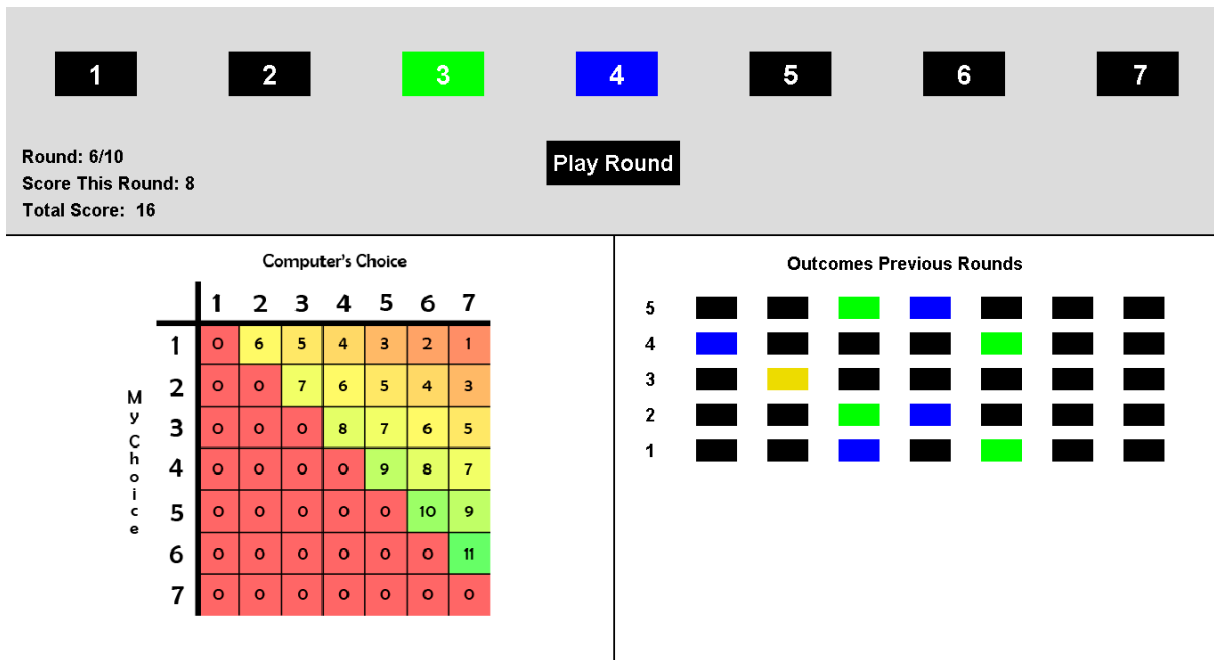
Figure 3.1: A screenshot of the reversed bidding game's interface. On top are the seven bidding options from which the participant could choose. Here, the opponent chose '4' and the player chose '3' in the previous round. The payoff matrix was also shown in the bottom left, and a history of the previous rounds in the bottom right.

opponent, and react to that accordingly. For example, when a ToM_1 agent has bid 6 in a round of the reversed bidding game, it can reason that if it were in its opponents position it would try to go for the undercut in the following round by bidding 5, and therefore bid 4 to outsmart its opponent.

Higher order theory of mind agents have the added ability of being able to reason about their opponent as if they were a theory of mind agent with a lower order than their own. A second-order theory of mind agent (ToM_2) for example, can reason that its opponent may be a ToM_1 agent. An example of how this reasoning would work in the reversed bidding game is when a ToM_2 agent has bid 6 in one round, it can reason that its opponent can be a ToM_1 agent that tries to go for the undercut by bidding 4 in the following round, and therefore the ToM_2 agent could bid 3 to counter this bid.

3.3. Materials & Procedure

Before the experiment began, the participants were guided to a lab with closed off cabins where the experiment took place. This was a quiet room at the University of Groningen that was specifically built for experiments, in order to avoid distractions. Once there, the participants were asked to read and sign an informed consent form, and enter their age, gender and education level into the laptop on which the game would be played. Before the game started, they were asked not to rush through the game, and to come to the person supervising the experiment if they had any questions or if anything was still unclear.

At the start of the game, the participants were shown instructions on how to play the game, and were informed that, after a short practice block, they would be playing against a different opponent with a different strategy for each block. These strategies are explained in detail in Section 3.3.1. The instructions also explained the interface of the game, which can be seen in Figure 3.1. Once the participants had read all instructions, they could press a button to start the practice game. During the game, they were shown seven bidding options, labelled one to seven. The participant had to choose one of them as their bid, without knowing what the computer would choose. Once the participants made their choice, they were shown the choice the computer made. This was done using colour-coding. The bidding option the participant chose became green, whereas the option the computer chose became blue. If both the computer and the participant chose the same bidding option, that option would become yellow. After each round a

3

visualization of the outcome, where it showed what both the player and the computer chose, was added to the game history. This history was displayed to the participant, to help them reason about what the computer would do next. The game existed of four separate blocks, each being ten rounds long.

The first of the four blocks in the experiment was the practice phase. In this part, the participants could practice for ten rounds, in order to make sure they understood how to play the game, and to allow them to familiarize themselves with the controls. The data from this block have not been used for the analysis. After the practice phase, a message was shown informing the participants that the practice phase was over, and the real part of the experiment - of which the data were recorded - would start. Furthermore, the message also stated that the participant should not continue with the experiment if any part was still unclear. The three blocks that followed were part of the experimental phase, which was used for the analysis.

3.3.1. Computer strategies

Participants played four blocks of the reversed bidding game, in which they faced a computer opponent with a different strategy for each block. This was done to see whether the participants used a different order of theory of mind when they had to play against different types of opponents.

The first block, which was the practice block, was against a *ToM2* agent. This was chosen so that the beginning of every game featured a reactive opponent, priming the participants not to look for static patterns. This is desirable, as the focus of this research was on the way the participants reasoned about what their opponent would do. When a participant would notice a clear pattern, they may no longer actively reason about their opponent, as they could simply follow the pattern mindlessly. Because the data from the practice block were not used for the research, it did not have to be in the same order within blocks, allowing for a reactive opponent such as a *ToM2* agent.

The opponents of the three blocks in the experimental phase had a pre-defined pattern they would follow. This allowed for more consistent data, as every participant was faced with the same pattern of responses from their opponent. Because of this, the choices of the computer in previous rounds influenced the reactions of the participants in the same way, as it was the same for all participants. To make sure that some effects were not caused by the order in which the different blocks were presented, the order of the three blocks the participants faced in the experimental phase was randomized.

During the three blocks of the experiment in the experimental phase, the different strategies the participants faced were positive, negative and random. The bids that the different strategies made during the game can be found in Table 3.1. The positive strategy generally bid higher than the two other strategies, but did try to defeat the player by luring them up, and then trying to undercut them. This can be seen in round 1-4, and again in round 5-7 and 8-10. The negative strategy bid a lot lower, trying to win by making a lot of low bids, hoping that the player continued to bid higher than itself. Lastly, the random strategy was a series of ten randomly generated bids, that was the same across participants. There was no further strategy behind it.

Table 3.1: Bids of the different computer strategies

	1	2	3	4	5	6	7	8	9	10
Positive	7	7	5	3	7	5	2	6	6	2
Negative	4	3	1	4	2	4	2	2	2	1
Random	4	4	2	7	6	6	1	3	7	1

3.4. Data

The obtained data were all the bids each participant made. This resulted in 29 arrays – one for each participant – each containing 30 bids. These data were used to create a distribution of the choices that all the participants made. For each round in a block, the bids that all participants made in that round were added to a table. This resulted in three 7 by 10 tables, filled with the choice distribution over the ten rounds. To determine what strategies were used by participants, the same type of bid distributions were also obtained for agents with different strategies, by making use of the simulations of the theory of mind agents. These agents were used to calculate the likelihoods for all bidding options for an agent with a certain
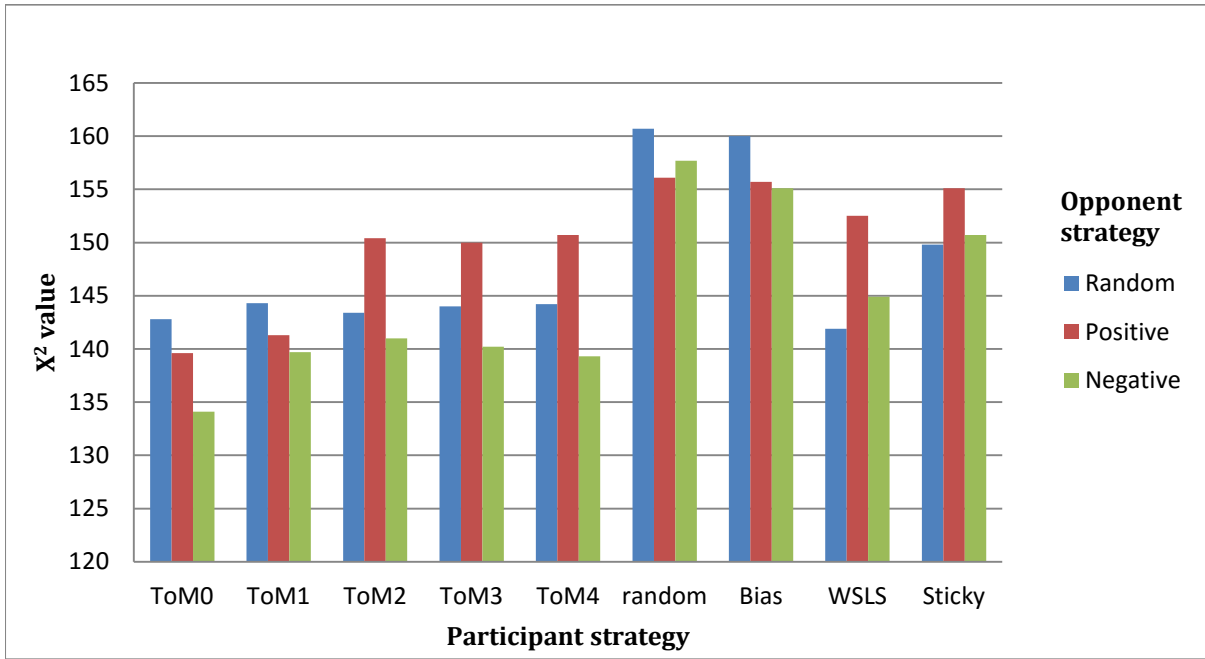
Figure 4.1: X² measurements of the different strategies for the reversed bidding game. The colour of the bars indicates the corresponding block

model or strategy. This was done for five different orders of theory of mind agents, ranging from zero-order theory of mind to fourth-order theory of mind, an agent with a random strategy that chose its bid randomly every round, and three more simple agent strategies: *Win-Stay-Lose-Shift* (*WSLS*), *Bias* and *Sticky*.

The *WSLS* strategy has proven useful before in the modelling of human behavior during a game of the prisoner's dilemma (Nowak & Sigmund, 1993). In the reversed bidding game *WSLS* chose its first bid randomly, and if it scored points with it, it would make the same bid again. If it didn't score points, it randomly chose another bid. The *Bias* strategy was used to see whether the participants favored a simple strategy with a more negative approach in general. The *Bias* strategy had a high probability of bidding one in the reversed bidding game, and would otherwise make a random bid between 2-7. Lastly, the *Sticky* strategy simulated a stubborn player, that would try to stay with its first bid. It randomly chose its first bid, and then had a high probability of repeating that bid in the following round. It was therefore likely to stay on the same bid.

These three simple strategies were added to the data analysis in order to see whether people resorted to one of these simple strategies, instead of choosing randomly or using one or more different orders of theory of mind to determine what bid they would make.

For each block, the distribution of the participant choices was compared with the expected distribution for each strategy, as they were calculated by the agents. Calculating the goodness of fit between the participant distribution and the distribution of a strategy showed how well that strategy could explain the participant data, and therefore give an insight into what strategy people most likely used during the experiment. To measure this goodness of fit, a chi-squared test was performed using Formula 2.1.

$$X^2 = \sum_{i,j} \frac{(ParticipantData_{ij} - AgentData_{ij})^2}{AgentData_{ij}} \qquad (2.1)$$

This provides a measure of distance between the two models, therefore a lower X² value indicated a better fit with a model. The X² values of the different orders of theory of mind agents were compared to each other to find which most accurately simulated the human data. These values were used to assess the accuracy of the theory of mind agents, and to determine what level of theory of mind was most likely used by the participants while playing the reversed bidding game.

4. Results

Figure 4.1 shows the resulting X² values for all strategies for the three different opponents. The three different colours indicate what block the
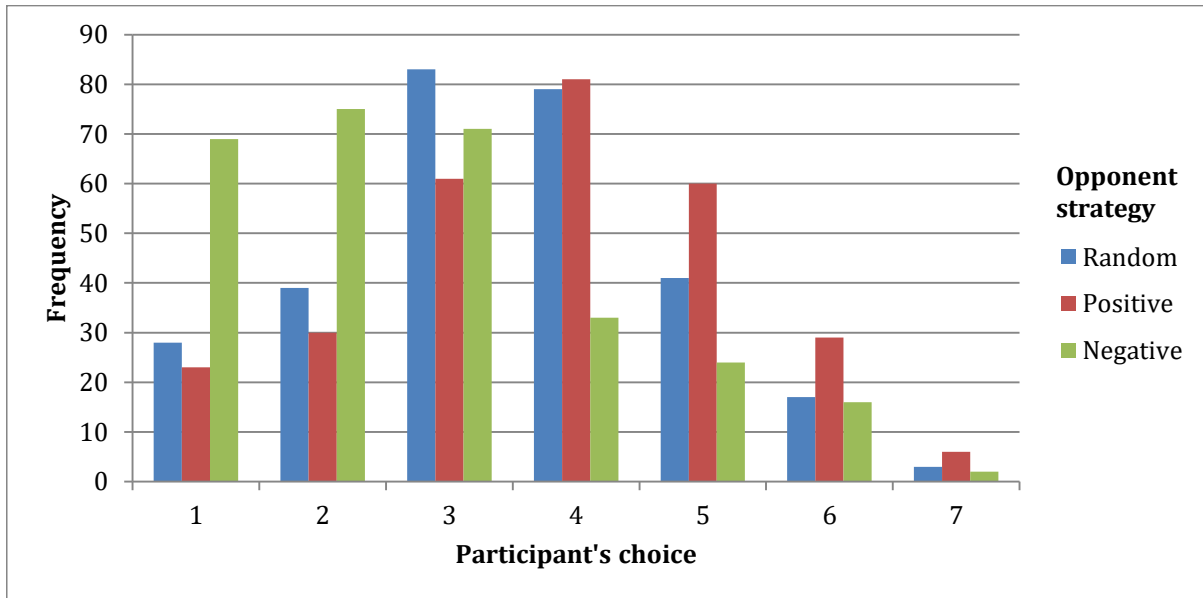
Figure 4.2: Distribution of all choices made by the participants per block. The y-axis notates the cumulative amount of times that the bidding option on the x-axis has been chosen per block. The colour of the bars indicates the corresponding block.

data were taken from. When the results of the different orders of theory of mind agents were compared, it showed that the zero-order agent had the best fit with the human data for all three opponent strategies. For the random opponent the values are relatively similar across all five orders of theory of mind agents, but for the negative and positive opponents the X^2 values did rise for higher orders of theory of mind. Up until the second order theory of mind agents the goodness of fit decreases, after which it stays roughly the same. These results suggest that people did not use theory of mind while playing the reversed bidding game.

It also shows that in all cases, the theory of mind agents had a better fit than the random model, suggesting that people did not pick their bids randomly. The rest of the simple strategies also had a worse fit than the theory of mind agents, apart from in one instance. Figure 4.1 shows that the *Win-Stay-Lose-Shift* strategy resulted in a low X^2 value, similar to *ToM0* and much lower than the other simple strategies during the random block. A possible explanation for this can be found in the way the theory of mind agents and the *WSLS* strategy work. The theory of mind agents make their decision based on the game's history, and predict the action of their opponent. Against a strategic player this can be advantageous, but against a player that selects their bid randomly every round this isn't beneficial. Because the history of the rounds

against a randomly playing opponent isn't very relevant, the simple *WSLS* strategy and the *ToM0* agent played very alike. This caused the fit of both strategies to also be similar, explaining why the simple *WSLS* strategy may have worked so well in this case. This theory is further strengthened by the fact that the fit for the positive and negative block for the *WSLS* strategy were a lot worse than the corresponding scores for the theory of mind models. As the positive and negative opponent were modelled to be probable response patterns, it shows in the results that the theory of mind models had a better fit than the *WSLS* strategy in these cases, as here the extended memory of previous rounds did aid the theory of mind models in playing more true to how the participants played the reversed bidding game.

Figure 4.1 also shows that for all theory of mind agents and for most other strategies the lowest X^2 values were achieved for the negative opponent. A possible explanation for this is that when people play against a defensive opponent, they start to play more defensively themselves as well. Once they realised that they were playing against an opponent that only made very low bids, they likely started making lower bids as well to get at least some points out of that block. It can be argued that this style of playing resulted in less unexpected moves, which therefore allowed the agents to make predictions that fit more accurately with the human data.

6

This notion is strengthened by the data that can be found in Figure 4.2, which shows the total amount of times each bidding option has been chosen per block across participants. The data from the negative block are more skewed towards the lower side of the bidding spectrum than the data from the positive block are skewed towards the upper spectrum. Another reason that may have caused the better fit for the negative block is the floor effect at the first bidding option. If a participant thought that the computer would choose 5 in the next round, they could choose 4, 3, 2 or 1 to try to score points, depending on how safe they wanted to play. However, if a participant thought that the computer would choose 2 in the next round, the only thing for them to choose to score any points was 1. This leaves the participants with fewer options to reasonably choose from, making it easier for the theory of mind agents to predict the next move of the participant.

Figure 4.2 unsurprisingly also shows that during the negative block the lower bidding options were chosen most frequently, whereas the participants bid higher during the positive block. The data from the random block are more evenly distributed over the bidding options. The graph does show that the participants on average favoured the lower bidding options, meaning that they often chose the safer option with a low return over a more risky option with a high return.

Figure 4.2 also shows that the bidding option 7 was bid only a few times. This is not unexpected, as in the reversed bidding game bidding 7 means that there is no way to score any points (see Table 2.1). However, it can be useful to try to get the opponent to make higher bids in the following rounds. It is therefore interesting that during the positive block this particular bidding option was chosen three times as often as during the negative block. This is unexpected, because if the participants had made use of theory of mind it would have been more logical for them to choose this option that is solely used to lure the opponent up during the negative block. This is the case because the most logical moment for a participant to make such a bid is when they thought that they could not score any points either way. This is possible when the opponent plays 1, which does not

happen during the positive block, whereas it does happen twice during the negative block. These results therefore fortify the notion that the participants did not make use of theory of mind while playing the reversed bidding game. These data suggest that the participants played very reactively, as they tended to follow the trend of their opponent, even if it was not the most beneficial thing to do.

5. Discussion & Conclusion

Earlier research on the topic of theory of mind has shown that people can gain an advantage by using theory of mind in competitive settings (De Weerd, Verbrugge & Verheij, 2013). As the goal of the experiment in this paper was to score as many points as possible in a competitive setting, it was expected that the participants would make use of higher order theory of mind while playing the game, as this could have given them an advantage in the amount of points they scored. Surprisingly, in this paper it was found that the participants showed behavior that matched most closely with the predictions of a zero-order theory of mind agent, suggesting that the participants did not make use of theory of mind.

During the experiment, human participants played a strategic closed bidding game called the reversed bidding game against a computer, where they could score the highest amount of points if they bid one bidding option lower than the computer opponent. This gave the participants an incentive to reason about the actions of their opponent. The selected bidding options were compared to agent based predictions of theory mind to analyze what type of theory of mind was most likely used by the participants. The surprising results were that the predictions of the zero-order theory of mind agent most accurately matched with the human data.

A possible reason for the lack of use of theory of mind by the participants can be the low number of game rounds during the experiment. Every block lasted for ten rounds. Other research with adults has shown that people reason at low orders of theory of mind, and only slowly adjust to opponents that do reason according to higher orders of theory of mind (Hedden & Zhang, 2002; Camerer et al., 2004; Wright & Leyton-Brown, 2010; Goodie et al., 2012). In games that

featured repeated rounds it is possible for participants to adjust their reasoning to use theory of mind to accurately predict the actions of the opponent, but this often takes the participants many rounds or trials to do (Zhang et al., 2012; Goodie et al., 2012; De Weerd et al., 2014; Devaine et al., 2014). Possibly, the participants did not have enough time to adjust to the different opponents in only ten rounds. One of the participants also said after the experiment that they did not really think very hard about their choices, as there was not enough rounds to figure out the opponent's strategy. Increasing the amount of rounds per block could result in different findings in future experiments.

Another possible explanation for these results is the type of game that was used. As most participants will not have ever done this type of bidding in reverse before, it may have been the case that it was an unintuitive game to play, as most people are more used to doing the opposite. This unintuitive nature of the game may have caused it to be harder to grasp for the participants, whom therefore may have resorted to simpler tactics to make decisions. Further research with a normal bidding game could give more clarity on this topic.

The agents used in this paper to compare the results to are also not perfect simulations of human behavior. A higher order theory of mind agent can reason about what another player will do, but it can't actively plan ahead as some players may do. For example, in a round of the reversed bidding game a higher order theory of mind agent will not bid 7 in order to get the opponent to bid higher in future rounds. This is a type of theory of mind reasoning that some of the participants may have used, but were not picked up by the theory of mind agents, as they are not capable of planning for future rounds.

Furthermore, the static opponents may also have influenced the choices of the participants. The static patterns, apart from the random pattern, were selected to give the participants the idea that they were playing against an active AI to avoid this type of influence on the results. If the participants however made an unexpected choice, the opponent could not realistically react to this. If, for example, a participant bid 7 twice in a row during the negative block to try to get the opponent to make higher bids, the opponent would not react to this and continue to make low bids. This may have made some participants notice that what they did had no effect on the bids of the computer, and therefore could have possibly changed the type of bids they made. In future research it would be interesting to see participants playing against a reactive agent during a different type of strategic reasoning game that is influenced less by the history of earlier rounds to make the research feasible.

References

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. The Quarterly Journal of Economics, 861-898.

de Weerd, H., Verbrugge, R., & Verheij, B. (2013b). How much does it help to know what she knows you know? An agent-based simulation study. Artificial Intelligence, 199, 67-92.

de Weerd, H., Verbrugge, R., & Verheij, B. (2014). Theory of mind in the Mod game: An agent-based model of strategic reasoning. In A. Herzig & E. Lorini (Eds.), Proceedings of the european conference on social intelligence (ECSI2014) (pp. 129–136).

Devaine, M., Hollard, G., & Daunizeau, J. (2014). The social Bayesian brain: does mentalizing make a difference when we learn?. PLoS Comput Biol, 10(12), e1003992.

Flobbe, L., Verbrugge, R., Hendriks, P., & Krämer, I. (2008). Children's application of theory of mind in reasoning and language. Journal of Logic, Language and Information, 17(4), 417-442.

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. Journal of Behavioral Decision Making, 25(1), 95-108.

Hedden, T., & Zhang, J. (2002). What do you think I think you think?: Strategic reasoning in matrix games. Cognition, 85(1), 1–36.

Meijering, B., van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2011). I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), Proceedings of the 33rd annual conference of the cognitive science society (pp. 2486–2491).

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature, 364(6432), 56-58.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1(4), 515–526.

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

Zhang, J., Hedden, T., & Chia, A. (2012). Perspective-Taking and Depth of Theory-of-Mind Reasoning in Sequential-Move Games. Cognitive science, 36(3), 560-573.