Theory of mind in the Mod game: An agent-based model of strategic reasoning

Harmen de Weerd, Rineke Verbrugge, Bart Verheij

Institute of Artificial Intelligence, University of Groningen

Abstract. When people engage in social interactions, they often rely on their *theory of mind*, their ability to reason about unobservable mental content of others such as beliefs, goals, and intentions. This ability allows them to both understand why others behave the way they do as well as predict future behaviour. People can also make use of *higher-order theory of mind* by applying theory of mind recursively, and reason about the way others make use of theory of mind such as in the sentence "Alice believes that Bob does not know about the surprise party". In this paper, we use agent-based models to describe human behaviour in an n-player extension of rock-paper-scissors called the Mod game. In previous work, we have shown how in similar competitive settings, the ability to make use of higher orders of theory of mind can be beneficial. We find that characteristic cyclic behaviour in the choices of participants that contradicts equilibrium predictions from classical game theory can be explained through the application of higher orders of theory of mind. Our results suggest that participants engage in higher orders of theory of mind reasoning in repeated play of the Mod game than previously reported in normal-form games and in repeated rock-paper-scissors games.

1 Introduction

People often make use of *theory of mind* [23] to explain and predict the behaviour of others. By reasoning explicitly about unobservable mental content such as beliefs, goals, and intentions, people are able to, for example, distinguish accidental from intentional behaviour. People can also make use of *higher-order theory of mind*, by using their own theory of mind ability to reason about the way others may use of their theory of mind. Second-order theory of mind allows people to form nested beliefs such as "Alice *believes* that Bob does not *know* about the surprise party", and use these beliefs to understand and predict the behaviour of Alice.

Empirical evidence shows that participants can make use of higher-order (i.e. at least second-order) theory of mind in tasks that require explicit reasoning about belief attributions of others [1, 22] as well as in strategic games [13, 18, 19]. However, there are limits to the depth of recursive theory of mind reasoning that people use [16, 21]. In particular, people are in general unable to explicitly use the infinite recursion needed to reason about *common knowledge* [11, 27]. As a result, people often fail to behave as predicted by equilibrium predictions of classical game theory.

In previous work, we presented an agent-based model of theory of mind [32] to show that agents benefit from the ability to make use of higher-order theory of mind in certain competitive settings. This model is closely related to hierarchical models of iterated reasoning in behavioural game theory, such as level-k reasoning [6, 25], cognitive hierarchies [5], quantal response equilibrium [17], and noisy introspection models [12]. In each of these models, the level of sophistication of agents is measured by the maximum number of steps of iterated reasoning the agent is capable of considering. In terms of theory of mind, one step of iterated reasoning approximately corresponds to zero-order theory of mind. Camerer at al. [5] estimate the distribution of the level of sophistication used by human participants over a range of normal-form games such as the p-beauty contest and the traveler's dilemma, and find an average of 1.5 steps of iterated reasoning. Part of the participants were found to use more than two steps of iterated reasoning, which suggests these participants may have been relying on second-order theory of mind. However, only few players were found to be well-described as higher-level agents [33].

Our model of theory of mind agents differs from previous models in that the behaviour of our agents changes based on the observed behaviour of others. Previous models typically assume that the most basic

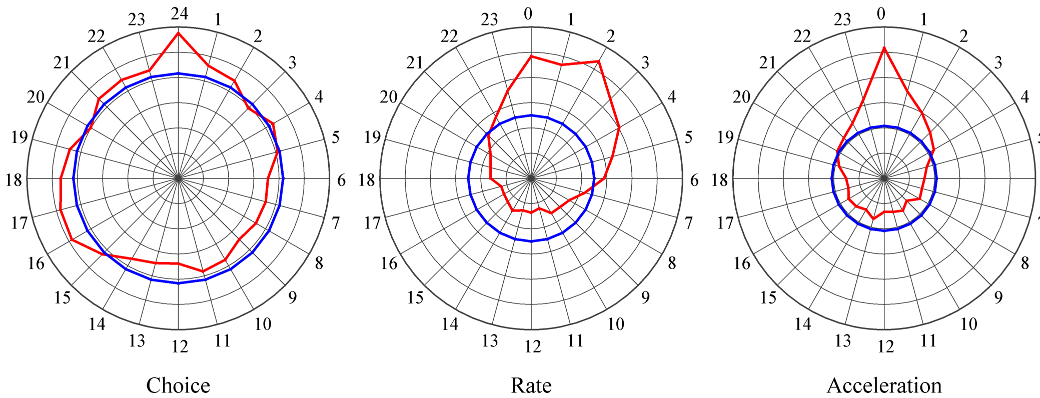


Fig. 1: Histograms over 24 choices, rates, and accelerations of human behaviour in the Mod game. In each graph, the blue curve shows the expected results from random behaviour, while the red curve shows the participant behaviour (reconstructed from [9]).

agent responds optimally under the assumption that other agents act according to a known non-strategic policy [5, 6, 12, 17, 25, 34]. Instead, our zero-order theory of mind agents attempt to learn the behaviour of others in repeated games through heuristics and associative learning. Our agent model has shown that in competitive settings such as repeated rock-paper-scissors games, second-order theory of mind can benefit agents greatly [32], while human behaviour in normal-form games suggests lower levels of recursive reasoning [5]. In the current paper, to determine whether human participants make use of higher-order theory of mind in these competitive settings, we describe human behaviour in the Mod game [10], an n-player extension of rock-paper-scissors, using our agent model. By comparing the behaviour of theory of mind agents with participant data from the Mod game, we can determine to what extent higher orders of theory of mind can account for the observed patterns of human behaviour.

In Section 2, we present a detailed description of the Mod game, as well as the way human participants play the game. Section 3 shows how theory of mind agents are implemented in this setting and how these agents play the Mod game. We simulate interactions between these theory of mind agents and compare the results to the behaviour of human participants. The results of these experiments are outlined in Section 4. Finally, Section 5 provides a discussion of the results as well as directions for future research.

2 Mod Game

The Mod game is an n-player generalization of rock-paper-scissors, introduced by Frey and Goldstone [10] as a way to reveal patterns in individual iterated reasoning strategies. In the Mod game, n participants simultaneously choose a number in the range $\{1, \ldots, m\}$, for both n and m greater than one. Participants gain one point for every other participant that has chosen the number that is exactly one lower than their own choice. For example, a participant that has chosen the number 4 gains a point for every participant that has chosen number 3. The only exception to this rule is that participants that have chosen number 1 gain one point for every participant that has chosen number m. The Mod game therefore has a structure similar to rock-paper-scissors, in which each action is dominated by some other action. In particular, the Mod game is a non-zero-sum version of rock-paper-scissors for $n = 2$ and $m = 3$.

The Mod game has a mixed-strategy Nash equilibrium in which each action is chosen with equal probability. When all players play according to this strategy, none of the players has an incentive to change his or her strategy. However, if a player deviates from the randomizing strategy, other players can take advantage from this regularity by switching their strategy as well. Interestingly, human participants are generally poor at generating random sequences [14, 24, 28]. This suggests that in groups of people, individuals can increase their score if they deviate from playing the Nash equilibrium strategy.

Indeed, participant behaviour in repeated Mod games deviates from the Nash equilibrium, as shown in Figure 1 (reconstructed from [9]). The figure shows the aggregate participant data (red curve) and the idealized randomizing behaviour (blue curve) over 100 rounds of play and $m = 24$. Participant choices appear to be approximately random, with a slight bias towards 24. However, *participant rates*, defined as the difference in choice between two subsequent rounds, shows a clear deviation from the Nash equilibrium in Figure 1. Participants are less likely to select numbers that are 7 to 21 ahead of their previous choice. Instead, they are most likely to choose a number that is 0 to 4 higher than their previous choice. *Participant acceleration*, which is defined as the change in participant rate, shows a similar effect. Figure 1 shows that participants tend to vary little in their rate. A participant who chose a number in the last round that was 2 higher than the number in the round before that is mostly likely to choose the number that is 2 higher than his choice in the previous round. However, Figure 1 shows that participants do vary their acceleration by a small amount.

Interestingly, this effect is not due to participants' poor performance on choosing random actions. When participants are given the option to let the computer select a randomly generated action, this option is used little [9]. This suggests that participants choose their actions based on their predictions of the behaviour of others rather than believing that the behaviour of others is unpredictable.

3 Theory of mind agents in the Mod game

In this section, we describe theory of mind agents that are able to play the Mid game outlined in Section 2. These agents are inspired by the theory of mind agents that we introduced in [32] to investigate the effectiveness of theory of mind in competitive settings. They engage in simulation-theory of mind [7, 20, 15] by taking the perspective of their opponents. An agent determines what he would do himself if he were facing the situation of his opponent, and attributes this thought process to his opponent to predict her behaviour. Each additional order of theory of mind allows the agent to generate an additional hypothesis about the way an opponent is playing the game. The task of a theory of mind agent is then to determine which hypothesis most accurately predicts the behaviour of his opponent.

The following subsections describe how agents of different orders of theory of mind play the n-player Mod Game [10].

3.1 Zero-order theory of mind

A zero-order theory of mind (ToM_0) agent has no theory of mind at all, and is therefore unable to attribute mental content to others. In particular, a ToM_0 agent cannot form the belief that his opponents are trying to obtain a high score. Instead, the ToM_0 agent forms zero-order beliefs about the collective actions of the agents playing the game. For each number, the ToM_0 agent specifies what he believes to be the likelihood that most players will select to play that number. Given these beliefs, the ToM_0 agent can select the number that he expects to maximize his score. For example, if a ToM_0 agent strongly believes that number 4 will be selected by most players, the agent should choose to play number 5 himself.

After every round, the ToM_0 agent updates his zero-order beliefs to reflect the actual outcome, such that the agent's new beliefs are constructed through a linear combination of his original beliefs and the newly observed situation. An agent-specific learning speed $\lambda \in [0, 1]$ determines the relative influence of the observation on the agent's beliefs. A ToM_0 agent with zero learning speed ($\lambda = 0$) does not update his beliefs at all. Such an agent selects the same action in every round. A ToM_1 agent with the maximal learning speed ($\lambda = 1$), on the other hand, completely replaces his zero-order beliefs after each observation, and forgets all information obtained from previous rounds. Such an agent considers the observed actions of the last round as the best predictor for the future.

The ToM_0 agent we describe here holds a simple model of agent behaviour. Although agents could model the behaviour of each individual opponent separately, our ToM_0 agent models the collective behaviour of others instead[1]. The ToM_0 agents also do not have any explicit memory. Although information obtained in the past is reflected in the agent's zero-order beliefs, the agent does not have an explicit representation of the past.

3.2 First-order theory of mind

Unlike the ToM_0 agent, a first-order theory of mind (ToM_1) agent reasons about the other's goals and therefore believes that his opponents may be trying to maximize their score as well. To predict the behaviour of his opponents, the ToM_1 agent attributes his own thought process to his opponents. Like the ToM_0 agent, the ToM_1 agent models a modal opponent rather than having a separate mental model for each individual opponent.

Following [32], the ToM_1 agent does not attempt to model the learning speed λ for his first-order model of opponent behaviour. Rather, the ToM_1 agent assumes that the modal opponent has the same learning speed as he has himself. This means that after a sufficient number of rounds, the ToM_1 agent predicts that most of the his opponents will play the action suggested by the agent's zero-order theory of mind. For example, for a ToM_1 agent that has strong zero-order beliefs that number 4 will be selected by most players, the agent's zero-order response would be to play number 5. Using first-order theory of mind, the ToM_1 agent attributes this though process to his opponents, and predicts that most of them will play number 5. The ToM_1 agent's first-order response would therefore be to play number 6.

Although the ToM_1 agent models his opponents as being able to use zero-order theory of mind, agents in our setup do not know the extent of the abilities of their opponents for certain. Rather, a ToM_1 agent has two models of opponent behaviour, one based on zero-order theory of mind and one on first-order theory of mind. Through repeated interaction, a ToM_1 agent learns which of his models best describes the behaviour of his opponents. Based on this information, a ToM_1 agent may therefore choose to play as if he were a ToM_0 agent, and ignore the predictions of his first-order theory of mind.

3.3 Higher orders of theory of mind

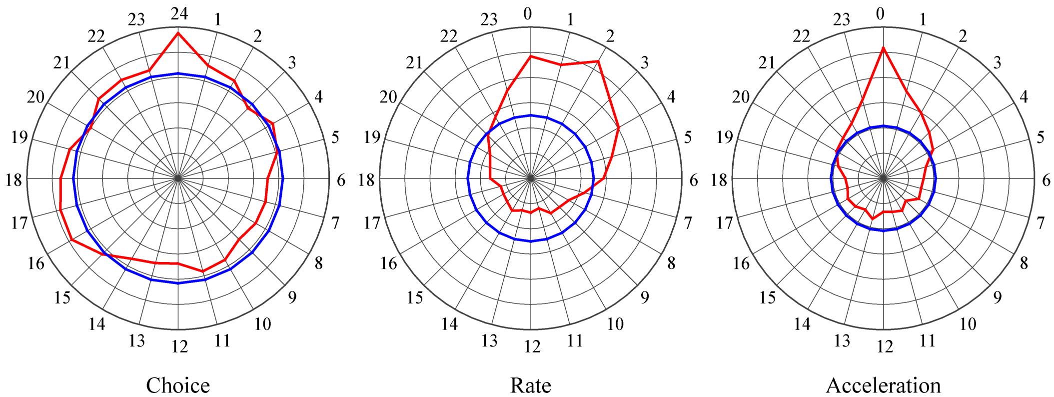
For each additional order of theory of mind k, an agent generates an additional prediction of opponent behaviour by attributing his own $(k-1)$st-order theory of mind thought process to his opponents. For example, a ToM_2 agent models the modal opponent as a ToM_1 agent, in addition to his zero-order and first-order theory of mind models of the modal opponent. As a result, a ToM_k agent has $k+1$ hypotheses for the action that will be chosen by most of his opponents with corresponding predictions. Based on the accuracy of these predictions, the ToM_k agent can therefore choose to behave according to $k+1$ patterns of behaviour.

Our agent model shows that in competitive settings such as repeated rock-paper-scissors games, second-order theory of mind can benefit agents greatly [32], while the additional benefit for fourth-order and even higher orders of theory of mind is limited [32]. In this paper, we therefore do not consider agents that are capable of theory of mind orders beyond the third.

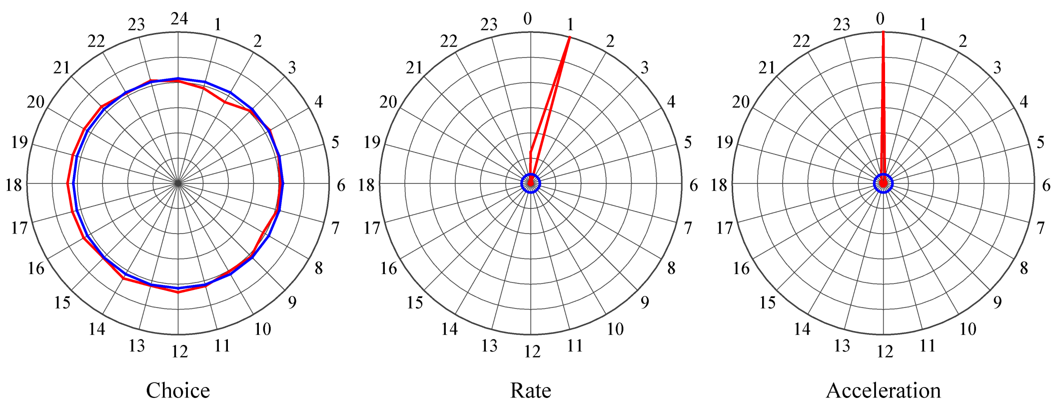
4 Simulation results

We simulate interactions of groups of the theory of mind agents of Section 3 in the Mod Game described in Section 2. To recreate the setting in [9, 10], agents repeatedly played the Mod Game for 100 rounds. In each condition, all agents had the same order of theory of mind, and were assigned randomized learning speeds. In [9], participants had the option to let the computer generate a random choice rather than to choose themselves, which was used in 9% of all choices. To simulate this, each action was assigned a 9% probability of being replaced with a randomly chosen action. That is, each agent had a 9% probability of selecting a random action rather than the one he believes to be the best action.

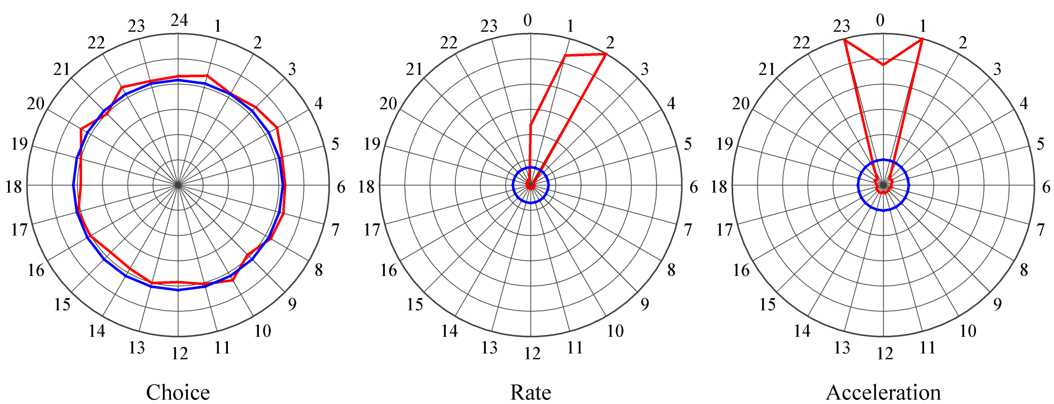
[1] Empirical evidence from weak-link coordination games suggests that participants only consider part of the data [8, 26]. However, evidence to the contrary also exists (see for example [3, 4]).



(a) Participant data (reconstructed from [10])

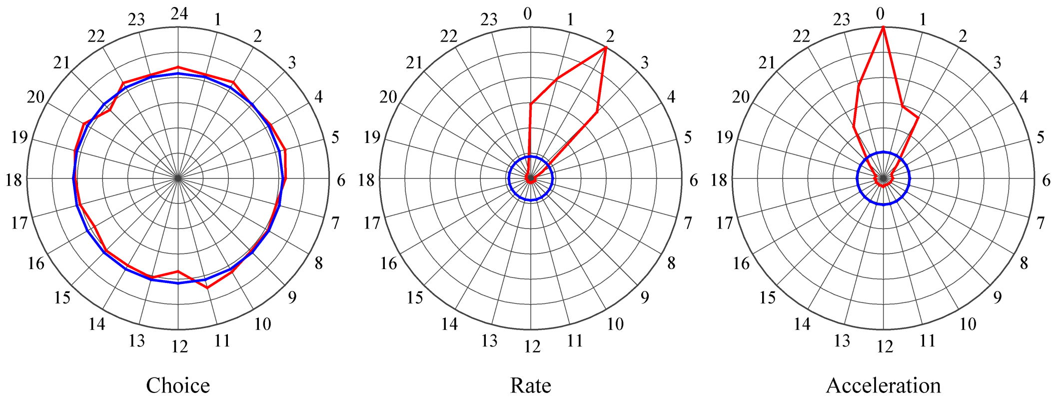


(b) Group of five ToM_0 agents

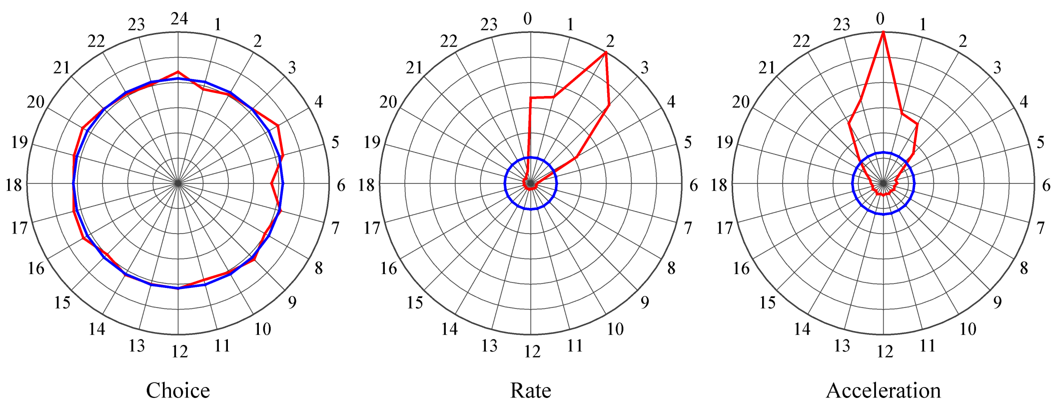


(c) Group of five ToM_1 agents

Fig. 2: Histograms over 24 choices, rates, and accelerations. In each graph, the blue curve shows the expected results from random behaviour, while the red curve shows the agent or participant behaviour.



(a) Group of five ToM_2 agents



(b) Group of five ToM_3 agents

Fig. 3: Histograms over 24 choices, rates, and accelerations. In each graph, the blue curve shows the expected results from random behaviour, while the red curve shows the agent or participant behaviour.

Figures 2 and 3 show histograms of the simulations that we have run with groups of agents that share the same ability to make use of theory of mind, as well as data from human participants reported in [9]. In each panel, the red curve shows the data obtained from participants or agent simulations, while the blue curve shows the histogram corresponding to the unique Nash equilibrium.

The first panel of each triple shows a histogram over the possible *choices*, aggregated over 100 rounds of play. Even though the theory of mind agents only have a 9% probability of choosing randomly, their choices superficially appear to resemble a random distribution, irrespective of the order of theory of mind of the agents.

The second panel of each triple shows the *rate* at which agents and participants change their choice. That is, these panels show the first difference in the choice of the participants and agents. The third panel shows the *acceleration* of agents and participants, which is the first difference in the rate of change of the choice of agents and participants.

Figure 2 shows that although ToM_0 and ToM_1 agents superficially appear to make random choices, these agents select their choices in a predictable pattern. Figure 2b shows that ToM_0 agents typically select the number that is one higher than the number that was chosen the most often in the previous round. That is, they select the number that would have won in the previous game. Since each ToM_0 agent acts the same way, ToM_0 agents typically have a rate of 1 and an acceleration of 0.

Figure 2c shows that ToM_1 agents display more variation in their behaviour. A ToM_1 agent typically selects the number that is either one or two higher than the number that was chosen most often in the previous round. The third panel of Figure 2c shows that, unlike ToM_0 agents, ToM_1 agents do not have zero acceleration. Rather, ToM_1 agents switch between choosing the number that would have won in the previous game and the number that is one higher. Although the rate data of ToM_1 agents fit participant data in Figure 2a better than ToM_0 agent data, acceleration data from ToM_1 agents differ from participant data. Whereas participants are most likely to keep their rate constant, ToM_1 agents are less likely to do so. Instead, ToM_1 agents are more likely to alternate between increasing and decreasing their rate.

For increasingly higher orders of theory of mind, agent behaviour shows increasingly more variation. Although the simplified agent architecture does not reach the variability seen in participant data, results from the higher-order theory of mind agent simulations show some interesting similarities to participant data. Similar to participants, higher-order theory of mind agents typically select numbers that are a little higher than the one they chose in the previous round. More precisely, ToM_2 agents tend to select a number that is up to 3 higher than their previous choice, while ToM_3 agents are also likely to select a number that is 4 higher than their previous choice. Note that this is true even though agents cannot remember their previous choices. Also, acceleration data from participants more closely resembles the acceleration data from higher-order theory of mind agents than that of ToM_0 and ToM_1 agents. Like participants, ToM_2 agents and ToM_3 agents display variation in their acceleration, but they are most likely to keep their rate constant. Interestingly, the higher-order agents also show the average negative acceleration reported in [9].

Our results show that the behaviour of theory of mind agents displays some interesting similarities to the behaviour of human participants in repeated Mod games. Moreover, human behaviour in the Mod game is closer to the behaviour of agents that make use of higher-order theory of mind than that of agents that are limited to zero-order or first-order theory of mind.

5 Discussion

Experimental studies show that people make use of higher-order theory of mind. In previous work, we have shown that in certain settings, agents can benefit from using higher orders of theory of mind. In cooperative settings, for example, first-order and second-order theory of mind can help to establish cooperation faster [30] even when a cooperative solution can be maintained without the use of theory of mind, while higher-order theory of mind can also provide an agent with a competitive advantage over others [32]. In this paper, we have compared these theory of mind agents with human behaviour in repeated play of the Mod game [10].

The Mod game has a mixed-strategy Nash equilibrium in which each action is played with equal probability. Contrary to predictions of classical game theory, participants do not appear to randomize their decision in every game. Rather, participant choices show cycles of varying speed [10]. Interestingly, our simulations with groups of theory of mind agents show that this kind of behaviour is a closer match to the behaviour of a group of agents capable of at least second-order theory of mind than to the behaviour of a group of agents that is more limited in their theory of mind abilities.

In previous work, the average depth of reasoning used by humans over a range of games was determined to be 1.5 steps [5], which corresponds approximately to first-order theory of mind. A recent study into human behaviour in repeated rock-paper-scissors games suggests that people similarly reason at low orders of theory of mind [2]. However, whereas participants in previous studies played unrepeated games or games that were repeated three times, participants played more repetitions in the Mod game. Our results suggest that over repeated plays of competitive games such as the Mod game, participants may increase their depth of reasoning and make use of higher orders of theory of mind.

For competitive settings such as the Mod game, our agent model suggests that reasoning at higher orders of theory of mind can benefit agents up to a certain point [32]. In particular, we found no advantage for the use of fourth-order theory of mind. Simulations in mixed-motive settings, in which both cooperative and competitive goals play a role, suggest that theory of mind also helps to stabilize mutually beneficial interactions [31]. In contrast to strictly competitive settings, reasoning using fourth-order theory of mind may be beneficial in negotiations [29]. It would be interesting to investigate whether human participants are capable of taking advantage of this benefit.

Acknowledgments

This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vici grant NWO 277-80-001, awarded to Rineke Verbrugge for the project 'Cognitive systems in interaction: Logical and computational models of higher-order social cognition'.

References

[1] Apperly, I.: Mindreaders: The Cognitive Basis of "Theory of Mind". Psychology Press, Hove, UK (2011)
[2] Batzilis, D., Jaffe, S., Levitt, S., List, J.A., Picel, J.: How facebook can deepen our understanding of behavior in strategic settings: Evidence from a million rock-paper-scissors games (2014), unpublished manuscript
[3] Berninghaus, S.K., Ehrhart, K.M.: Coordination and information: Recent experimental evidence. Economics Letters 73(3), 345–351 (2001)
[4] Brandts, J., Cooper, D.J.: Observability and overcoming coordination failure in organizations: An experimental study. Experimental Economics 9(4), 407–423 (2006)
[5] Camerer, C., Ho, T., Chong, J.: A cognitive hierarchy model of games. Quarterly Journal of Economics 119(3), 861–898 (2004)
[6] Costa-Gomes, M., Crawford, V.P., Broseta, B.: Cognition and behavior in normal-form games: An experimental study. Econometrica 69(5), 1193–1235 (2001)
[7] Davies, M.: The mental simulation debate. Philosophical Issues 5, 189–218 (1994)
[8] Devetag, G.: Precedent transfer in coordination games: An experiment. Economics Letters 89(2), 227–232 (2005)
[9] Frey, S.: Complex collective dynamics in human higher-level reasoning: A study over multiple methods. Ph.D. thesis, Indiana University (2013)
[10] Frey, S., Goldstone, R.L.: Cyclic game dynamics driven by iterated reasoning. PLoS ONE 8(2), e56416 (2013)
[11] Gintis, H.: The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences. Princeton University Press, Princeton (NJ) (2009)
[12] Goeree, J.K., Holt, C.A.: A model of noisy introspection. Games and Economic Behavior 46(2), 365–382 (2004)
[13] Hedden, T., Zhang, J.: What do you think I think you think?: Strategic reasoning in matrix games. Cognition 85(1), 1–36 (2002)
[14] Herbranson, W.T., Schroeder, J.: Are birds smarter than mathematicians? Pigeons (*Columba livia*) perform optimally on a version of the Monty Hall Dilemma. Journal of Comparative Psychology 124(1), 1–13 (2010)
[15] Hurley, S.: The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. Behavioral and Brain Sciences 31(01), 1–22 (2008)
[16] Keysar, B., Lin, S., Barr, D.: Limits on theory of mind use in adults. Cognition 89(1), 25–41 (2003)
[17] McKelvey, R., Palfrey, T.: Quantal response equilibria for normal form games. Games and Economic Behavior 10(1), 6–38 (1995)
[18] Meijering, B., van Rijn, H., Taatgen, N., Verbrugge, R.: I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In: Proceedings of the 33nd Annual Conference of the Cognitive Science Society. pp. 2486–2491. Cognitive Science Society (2011)
[19] Meijering, B., Van Maanen, L., Van Rijn, H., Verbrugge, R.: 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. Cognitive Science Society (2010)
[20] Nichols, S., Stich, S.: Mindreading: An Integrated Account of Pretence, Self-awareness, and Understanding Other Minds. Oxford University Press, USA (2003)
[21] Ohtsubo, Y., Rapoport, A.: Depth of reasoning in strategic form games. The Journal of Socio-Economics 35(1), 31–47 (2006)
[22] Perner, J., Wimmer, H.: "John thinks that Mary thinks that...". Attribution of second-order beliefs by 5 to 10 year old children. Journal of Experimental Child Psychology 39(3), 437–71 (1985)
[23] Premack, D., Woodruff, G.: Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences 1(04), 515–526 (1978)

[24] Rapoport, A., Budescu, D.: Randomization in individual choice behavior. Psychological Review 104(3), 603 (1997)

[25] Stahl, D., Wilson, P.: On players' models of other players: Theory and experimental evidence. Games and Economic Behavior 10(1), 218–254 (1995)

[26] Van Huyck, J.B., Battalio, R.C., Beil, R.O.: Tacit coordination games, strategic uncertainty, and coordination failure. The American Economic Review pp. 234–248 (1990)

[27] Verbrugge, R.: Logic and social cognition: The facts matter, and so do computational models. Journal of Philosophical Logic 38, 649–680 (2009)

[28] Wagenaar, W.: Generation of random sequences by human subjects: A critical survey of literature. Psychological Bulletin 77(1), 65–72 (1972)

[29] de Weerd, H., Verbrugge, R., Verheij, B.: Agent-based models for higher-order theory of mind. In: Advances in Social Simulation, Proceedings of the 9th Conference of the European Social Simulation Association, vol. 229, pp. 213–224. Springer, Berlin (2014)

[30] de Weerd, H., Verbrugge, R., Verheij, B.: Higher-order theory of mind in Tacit Communication Game. In: Biologically Inspired Cognitive Architectures 2014: Proceedings of the Fifth Annual Meeting of the BICA Society (2014), to appear

[31] de Weerd, H., Verbrugge, R., Verheij, B.: Higher-order theory of mind in negotiations under incomplete information. In: Boella, G., Elkind, E., Savarimuthu, B., Dignum, F., Purvis, M. (eds.) PRIMA 2013: Principles and Practice of Multi-Agent Systems, pp. 101–116. Lecture Notes in Computer Science, Springer (2013)

[32] de Weerd, H., Verbrugge, R., Verheij, B.: How much does it help to know what she knows you know? An agent-based simulation study. Artificial Intelligence 199–200, 67–92 (2013)

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

[34] Yoshida, W., Dolan, R.J., Friston, K.J.: Game theory of mind. PLoS Computational Biology 4(12), e1000254 (2008)