

Training the use of theory of mind using artificial agents

Kim Veltman · Harmen de Weerd · Rineke Verbrugge

the date of receipt and acceptance should be inserted later

Abstract When engaging in social interaction, people rely on their ability to reason about unobservable mental content of others, which includes goals, intentions, and beliefs. This so-called theory of mind ability allows them to more easily understand, predict, and influence the behavior of others. People even use their theory of mind to reason about the theory of mind of others, which allows them to understand sentences like ‘Alice believes that Bob does not know about the surprise party’. But while the use of higher orders of theory of mind is apparent in many social interactions, empirical evidence so far suggests that people do not use this ability spontaneously when playing strategic games, even when doing so would be highly beneficial. In this paper, we attempt to encourage participants to engage in higher-order theory of mind reasoning by letting them play a game against computational agents. Since previous research suggests that competitive games may encourage the use of theory of mind, we investigate a particular competitive game, the Mod game, which can be seen as a much larger variant of the well-known rock-paper-scissors game. By using a combination of computational agents and Bayesian model selection, we simultaneously determine to what extent people make use of higher-order theory of mind reasoning, as well as to what extent computational agents can encourage the use of higher-order theory of mind in their human opponents.

Our results show that participants who play the Mod game against computational theory of mind agents adjust their level of theory of mind reasoning to that of their computer opponent. Earlier experiments with other strategic games show that participants only engage in low orders of theory of mind reasoning. Surprisingly, we find that participants who knowingly play against second- and third-order theory of mind agents apply up to fourth-order theory of mind themselves, and achieve higher scores as a result.

1 Introduction

Many social skills vitally depend on the ability of the person to reason about others as goal-oriented agents with their own beliefs, goals, and intentions, an important part of social cognition. This ability to reason explicitly about unobservable mental content of others, known as *theory of mind* [20], has been associated with pro-social behavior [13], social competences [15], and negotiation skills [31], as well as in producing and interpreting prosody [4] and nonverbal communication through body language and gestures [18, 19]. But while adults show impressive theory of mind abilities in some experiments that rely on communication, people are typically slow to take advantage of their theory of mind ability in strategic settings [11, 3, 32, 10]. In this paper, we explore the possible use of artificial theory of mind agents in quantifying and encouraging the use of theory of mind.

People do not only use their theory of mind to reason about goals, desires, and beliefs concerning world facts of others. Rather, people are able to use their theory of mind ability recursively, and reason about the way others make use of theory of mind. For example,

K. Veltman · H. de Weerd · R. Verbrugge
Department of Artificial Intelligence, Bernoulli Institute, University of Groningen

H. de Weerd
Research group User-Centered Design, Hanze University of Applied Sciences

people make use of *second-order theory of mind* to understand a sentence such as “Alice *knows* that Bob *knows* that Carol is throwing him a birthday party”, by reasoning about what Alice knows about what Bob knows. In experimental story comprehension tasks, adults show their impressive ability for recursive theory of mind ability, and score much better than chance on questions that explicitly involve fourth-order theory of mind reasoning [14, 23].

In this article, we explore the potential of artificial agents to train the social skills of human participants. In particular, we are interested in encouraging people to spontaneously engage in higher-order theory of mind reasoning, which is the basis for a variety of social skills. For example, it has been shown that when children with autism spectrum disorder are trained with theory of mind tasks, their social skills improve [1]. To accomplish our goal, we formulate the following two research questions.

1. To what extent do human participants make use of theory of mind reasoning when playing against artificial agents?
2. To what extent can interacting with artificial agents encourage human participants to make use of higher-order theory of mind reasoning?

To this end, we let participants play the Mod game [9] against artificial theory of mind agents, and estimate their level of theory of mind reasoning using random-effects Bayesian model selection [22].

Studies involving artificial agents with the ability to reason about the beliefs and goals of others show that higher-order theory of mind reasoning can be particularly effective in competitive settings [9, 6, 7]. Our previous research on the matching pennies game, however, shows that in this simple competitive game, many people rely on simpler, behavior-based strategies when engaging with artificial agents [27]. A possible cause for the lack of theory of mind reasoning was that due to the limited number of possible actions, it is difficult to distinguish between strategies. For this reason, we consider an extension of the matching pennies game with more possible actions, known as the Mod game. While the Mod game has a structure that is very similar to that of matching pennies, the larger number of possible actions should make it easier for participants to distinguish between strategies. This may help participants to reason about the goals and beliefs of their opponent, and encourage them to make use of higher orders of theory of mind. Moreover, our results from agent simulations in variants of the Mod game suggest that both first-order and second-order theory of mind can greatly benefit players, while the use of orders of theory of mind

beyond the second hardly provides additional benefits [30, 28].

In this Mod game setting, we let human participants play against virtual agents that we previously developed to determine the effectiveness of making use of increasingly higher orders of theory of mind [29]. By making use of artificial agents, we can precisely control and monitor the mental content of the opponents that participants face, including their application of theory of mind. This allows us to analyze and diagnose participant data from a more controlled setting, as well as ensure that participants play against an opponent that reasons at a particular level of theory of mind. The agents therefore provide us with a tool to diagnose human behavior (cf. research question 1). An additional benefit of using virtual agents is that by letting human participants train with virtual agents that are programmed to reason in a certain fashion, we can potentially expose participants to a level of theory of mind reasoning that would stimulate them to improve their own reasoning (cf. research question 2). For example, a participant who plays against a second-order theory of mind agent might recognize the reasoning strategy and apply third-order theory of mind to outsmart this tough opponent. Unlike human opponents, the agent is consistent in its use of theory of mind, which may make it easier for the participant to recognize the benefit of higher-order theory of mind reasoning.

As far as we know, explicit higher-order theory of mind training has not yet been part of virtual training agents for social skills, although some authors mention its possible usefulness for contexts in which deception plays a role [2, 16].

The remainder of this paper is structured as follows. Section 2 introduces and analyzes the Mod game, while Section 3 describes a range of strategies that agents and humans could use in this setting. In Section 4, we describe a method to gauge agents’ and participants’ reasoning strategies from their behavior known as random-effects Bayesian model selection. Section 5 delineates our experiment in which human participants played the Mod game against agents that used different orders of theory of mind. The results of this experiment are presented in Section 6. Section 7 concludes the article and describes how virtual agents can indeed be used to support people in using higher orders of theory of mind in a competitive game such as the Mod game. A preliminary version of this research was presented at BNAIC 2017 [24].

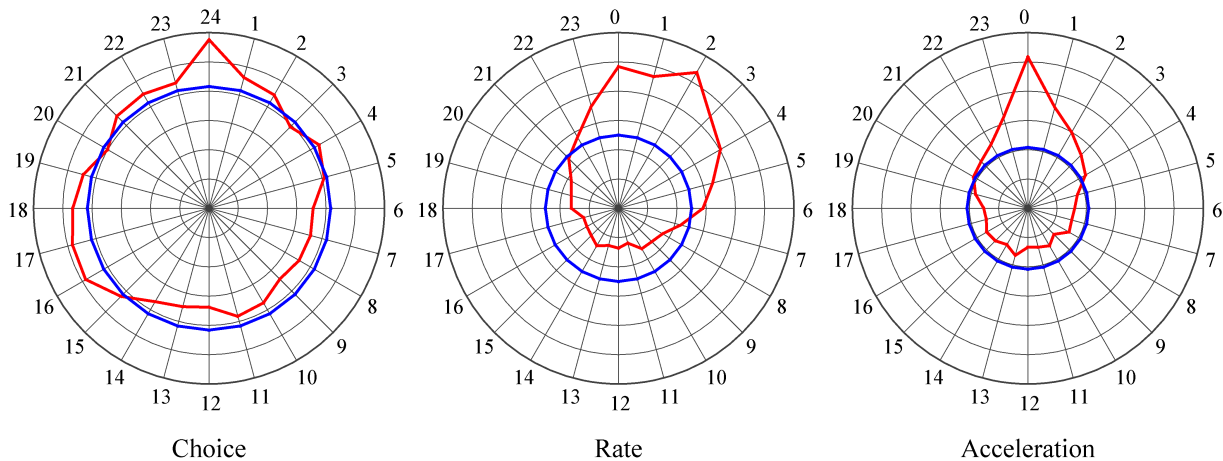


Fig. 1: Histograms over 24 choices, rates, and accelerations of human behaviour in the Mod game. In each graph, the blue curve shows the expected results from random behaviour, while the red curve shows the participant behaviour (reconstructed from [8]).

2 Mod Game

The Mod game is an n -player generalization of rock-paper-scissors, introduced by Frey and Goldstone [9] as a way to reveal patterns in individual theory of mind strategies. The Mod game is played by n players, who simultaneously choose a number in the range $\{1, \dots, m\}$, with $m > n > 1$. For every opponent that has chosen the number that is exactly one lower than their own choice, players gain one point. For example, a player that has chosen the number 4 gains a point for every opponent that has chosen number 3. The only exception to this rule is that players that have chosen number 1 gain one point for every player that has chosen number m . That is, the name ‘mod’ game refers to the goal of players to choose the number that is ‘+1 mod m ’ the number of their opponent(s). In our experiment, we visualize the rules of the game to human participants by arranging actions in a circle (see Figure 2 for $m = 24$).

Each action in the Mod game is dominated by some other action, similar to games such as rock-paper-scissors. In fact, the Mod game is equivalent to a non-zero-sum version of rock-paper-scissors for $n = 2$ and $m = 3$. Similar to rock-paper-scissors, the Mod game has a mixed-strategy Nash equilibrium in which each action is chosen with equal probability. That is, when all players play according to this randomizing strategy, none of the players has an incentive to change his or her strategy.

However, it is unlikely that a group of human participants would play according to this Nash equilibrium. Experimental evidence has shown that human participants are generally poor at generating random

sequences [12,21,26]. This suggests that in groups of people, it is likely that at least one person will deviate from playing the Nash equilibrium strategy. But if some player i deviates from this randomizing strategy, then all other players have an incentive to deviate from random play as well. After all, each player can increase their expected payoff by adjusting their strategy to take advantage of the predictability in the behavior of player i . With human players, social skills therefore play a role in the Mod game, because the person who can predict the beliefs and actions of others most accurately achieves the highest score.

Participant behavior in repeated Mod games indeed deviates from the Nash equilibrium, as depicted in Figure 1 (reconstructed from [8]). The figure shows experimental data for the Mod game with 24 actions, including the proportion of times a given number was chosen (red line in left graph) and the idealized randomizing behavior (blue line) across participants over 100 rounds of play. Participant choices (red line in left-most graph) appear to be approximately random, with a slight bias towards 24. However, a clear deviation from the Nash equilibrium is shown when the previous choice of a participant is considered. The middle graph in Figure 1 depicts *participant rates*, which is defined as the difference in choice between two subsequent rounds. As the figure shows, participants (red line in the middle graph) are most likely to choose a number that is 0 to 4 higher than their previous choice, while they are very unlikely to select numbers that are 7 to 21 ahead of their previous choice. If participants were to play according to the Nash equilibrium, however, each rate should be equally

likely (blue line). *Participant acceleration*, which is defined as the change in participant rate, shows a similar effect. Figure 1 (red line in right-most graph) shows that participants tend to vary little in their rate. That is, 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 in the current round than his choice in the previous round. In addition, Figure 1 also shows that participants that vary their acceleration do so by a small amount. Again, Nash equilibrium play would result in each acceleration being equally likely (blue line).

Frey and Goldstone [8] show that these deviations from the Nash equilibrium strategy are 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 [8]. This suggests that participants believe that they can accurately predict the actions of others, and act on this belief by choosing a number rather than going for the randomizing option. That is, participants may rely on social cognition when playing this game.

In our current experiment, participants play a specific variant of the Mod game with $m = 24$ actions and $n = 2$ players. For the remainder of the paper, we will only consider this specific variant of the Mod game.

3 Strategies in the Mod game

The Mod game, as outlined in Section 2, can be played using a variety of strategies. In this section, we describe a number of these strategies. These include strategies based on the use of theory of mind, as well as simple behavior-based strategies. Table 1 shows Mod game strategies we consider in this research. In addition to the theory of mind strategies, which are the main focus of our research, we consider several simpler, behavior-based strategies that rely purely on the actions observed in the previous round of play, as suggested by the results depicted in Figure 1.

In this section, we describe these strategies in detail. To avoid confusion, in the remainder, we will refer to focal agents as if they were male, while we will refer to their opponents as if they were female.

3.1 Behavior-based strategies

While participants may benefit from reasoning about the beliefs and goals of others while playing the Mod game, the game can be played without relying on such

Table 1: We consider eight possible strategies for playing the Mod game, including four behavior-based strategies and four theory of mind strategies.

Strategies	Behavior-based	ToM
Other regarding	✓	
Self regarding	✓	
Win-Stay-Lose-Shift	✓	
ToM_0	✓	
ToM_1		✓
ToM_2		✓
ToM_3		✓
ToM_4		✓

strategies. In our Bayesian RFX-BMS analysis of participant behavior, we therefore consider a number of behavior-based strategies. A player that uses a behavior-based strategy responds to actions observed in previous rounds of play only.

3.1.1 Self-regarding strategy

An agent that follows a *self-regarding* strategy ignores the actions of the opponent, and decides what action to take based on what action he has performed in the previous round. The self-regarding strategy depends on two free parameters. The drift parameter k determines the change in action in every round, so that an agent that follows a self-regarding strategy with drift k tends to choose the number that is k higher (modulo 24) than the action he performed in the previous round.

The choice probability p determines the strength of this self-regarding tendency. For example, an agent that follows a self-regarding strategy with $k = 2$ selects the action that is 2 higher than his previous choice with probability p , while each other action has a probability $\frac{1}{23}(1 - p)$ of being selected.

3.1.2 Other-regarding strategy

The *other-regarding* strategy is similar to the self-regarding strategy, except that an agent that follows the other-regarding strategy reacts to the previous action of his opponent rather than his own previous action. Like the self-regarding strategy, the other-regarding strategy relies on a drift parameter k and choice probability p . An agent that follows an other-regarding strategy with drift k selects the action that is exactly k higher (modulo 24) than his opponent's action in the previous round with probability p . Each other action is selected with probability $\frac{1}{23}(1 - p)$.

Note that if an agent plays according to an other-regarding strategy with $k = 1$, the agent tends to play the action that would have won in the previous round.

3.1.3 Win-Stay, Lose-Shift (WSLS) strategy

An agent that follows the *win-stay, lose-shift (WSLS)* strategy bases his current decision on the outcome of the previous round. If the agent won the previous round, he will repeat his previously chosen action with probability p , while each of the other 23 actions is selected with probability $\frac{1}{23}(1-p)$. However, if the agent did not win the previous round, he will repeat his previously chosen action with probability $1-p$, while each of the other 23 actions is selected with probability $p/23$. The single parameter p is a free parameter.

3.2 Theory of mind strategies

In addition to the behavior-based strategies described above, our analysis includes strategies that are based on taking advantage of social cognition while playing the Mod game. These strategies are inspired by the theory of mind agents that we introduced to investigate the effectiveness of theory of mind in competitive settings [30]. A theory of mind agent can take the perspective of other agents, a skill that lies at the root of many social skills. By determining what the agent would do himself if he were facing the situation of an opponent and attributing this thought process to that opponent, the theory of mind agent can formulate a prediction of the behavior of other players. Additional orders of theory of mind allow the agent to generate additional hypotheses of opponent behavior. The task of a theory of mind agent is then to determine which hypothesis yields the most accurate predictions. Below, we briefly describe these agents. A full mathematical model of these agents can be found in [30].

3.2.1 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 consider his opponent as a goal-directed agent who is trying to obtain a high score for herself. Instead, the ToM_0 agent forms zero-order beliefs about the actions the opponent will play in future rounds of the game based on her behavior in the past.

In our agent model, a ToM_0 agent forms beliefs $b^{(0)}$ about the actions of the opponent. For each number $i = 1, \dots, 24$, the ToM_0 agent has a belief $b^{(0)}(i)$ that represents what he believes to be the likelihood that his opponent will select to play that number. Given these beliefs, the ToM_0 agent can calculate the expected value $EV^{(0)}(i; b^{(0)})$ of choosing number i . Note that in the case of the Mod game, the expected value of choosing

number i is the belief that the opponent will choose the action $i - 1$ (modulo 24). That is,

$$EV^{(0)}(i; b^{(0)}) = n \cdot b^{(0)}\left((i-1) \bmod 24\right). \quad (1)$$

The ToM_0 agent acts on these beliefs by choosing the number that maximizes his score. For example, if a ToM_0 agent strongly believes that number 4 will be selected by his opponent, 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. An agent-specific learning speed $\lambda \in [0, 1]$ determines the relative influence of the current observation on the agent's beliefs. For example, 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_0 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 future¹.

To account for small deviations between participant choices and the ToM_0 agent strategy, we make use of the so-called ‘softmax’ probabilistic policy [5, 27]. That is, in addition to the learning speed λ , the ToM_0 agent strategy has an additional parameter β that controls the magnitude of behavioral noise, so that the probability that a ToM_0 agent chooses number i is

$$\begin{aligned} P(A = i) &= s\left(\frac{EV^{(0)}(i; b^{(0)})}{\beta}\right) \\ &= \frac{\exp(EV^{(0)}(i; b^{(0)})/\beta)}{\sum_j \exp(EV^{(0)}(j; b^{(0)})/\beta)}. \end{aligned} \quad (2)$$

As a result, the ToM_0 strategy has two free parameters: the behavioral noise parameter β and the learning speed λ .

3.3 First-order theory of mind

Unlike the ToM_0 agent, a first-order theory of mind (ToM_1) agent is capable of reasoning about the goals of others, and believes that his opponent may be trying to maximize her score. To predict the behavior of his opponent, the ToM_1 agent attributes his own thought process to her. A ToM_1 agent therefore considers the possibility that his opponent is a goal-directed agent like he is, and that while the agent reacts to the actions of his opponent, the opponent is reacting to the actions of the agent.

¹ Note that a ToM_0 agent with learning speed $\lambda = 1$ behaves identically to an other-regarding agent with $k = 1$.

Following our theory of mind models of [30], the ToM_1 agent does not attempt to model the learning speed λ for his first-order model of opponent behavior. Instead, the ToM_1 agent assumes that his opponent has the same learning speed as he has himself.

Although the ToM_1 agent models his opponent as being able to use zero-order theory of mind, agents in our setup do not know the extent of the abilities of their opponent for certain. Rather, a ToM_1 agent has two models of opponent behavior, one based on zero-order theory of mind and one on first-order theory of mind. Each of these models makes a prediction of the opponent's behavior. In addition, the ToM_1 agent has a confidence parameter c_1 ($0 \leq c_1 \leq 1$) that determines to what extent the agent's behavior is determined by first-order theory of mind reasoning. After each observation of the opponent's action a_j , this confidence is updated according to the following rule:

$$c_1 := \begin{cases} (1 - \lambda) \cdot c_1 + \lambda & \text{if } a_o = \hat{a}_o^{(1)}, \\ (1 - \lambda) \cdot c_1 & \text{otherwise,} \end{cases}$$

where $\hat{a}_o^{(1)}$ is the first-order theory of mind prediction of the action of the opponent.

Through repeated interaction, a ToM_1 agent learns which of his models best describes the behavior of his opponent. 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.

To account for behavioral noise, we apply a softmax policy to the actions prescribed by the ToM_1 agent strategy as well. However, this policy is applied only to the actions of the ToM_1 agent. The agent does not apply this policy in its model of opponent behavior. Although the ToM_1 strategy is more complex than the ToM_0 strategy, both strategies rely on the same two free parameters, namely behavioral noise β and learning speed λ . That is, the ToM_1 strategy does not introduce any additional free parameters.

3.4 Higher orders of theory of mind

For each additional order of theory of mind k , an agent generates an additional prediction of opponent behavior, by attributing his own $(k-1)$ st-order theory of mind thought process to his opponent. For example, a ToM_2 agent models his opponents as ToM_1 agents, in addition to his zero-order and first-order theory of mind models of opponent behavior. As a result, a ToM_k agent has $k + 1$ hypotheses for the action that will be chosen by his opponent, 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 behavior.

As described for the ToM_1 strategy, we apply a softmax policy to account for behavioral noise, which is applied only to the action that a focal ToM_k agent performs. In particular, the softmax policy is not applied to any of the ToM_k agent's models of opponent behavior. Also, while each additional order of theory of mind provides an agent with an additional prediction of opponent behavior, no additional parameters are introduced. That is, each theory of mind strategy is defined by its behavioral noise parameter β and learning speed λ .

4 Random-effects Bayesian model selection

In this paper, we attempt to encourage participants in their use of social cognition through interactions with artificial theory of mind agents. To determine what level of theory of mind reasoning a participant is engaging in at different points throughout the experiment, we make use of a technique known as group-level random-effects Bayesian model selection (RFX-BMS), introduced by Stephan and colleagues [22]. Whereas fixed-effects Bayesian model selection assumes that the actions of all participants can be best described by a single strategy, random-effects Bayesian model selection allows for individual differences in strategy selection. Strategies are treated as random effects that occur with an unknown but fixed probability in the population. A group of participants represents a random sample drawn from these strategies.

Random-effects Bayesian model selection estimates what distribution of strategies best fits the experimental data. Each strategy s generates pieces of evidence $p(y_i|s)$ representing the probability that choosing actions according to strategy s will result in some observed data y_i of participant i . Fixed-effects Bayesian model selection aims to identify the strategy s that has maximal evidence $\prod_i p(y_i|s)$ across all participants i . That is, fixed-effects Bayesian model selection assumes that there is a single strategy that explains the behavior of all participants.

In contrast, random-effects Bayesian model selection aims to identify the distribution of strategies in the population. That is, it aims to identify the relative frequencies f_s of strategies s with $\sum_s f_s = 1$ so that evidence $f_s \cdot \prod_i p(y_i|s)$ is maximized.

To determine to what extent a participant makes use of theory of mind while playing the Mod game, we compare the observed behavior y_i of each participant i with the predicted behavior of computational agents

following the strategies described in Section 3. That is, the model evidence $p(y_i|s)$ generated by a given strategy model s is the probability that following strategy s will result in the same behavior y_i as participant i .

The combination of our theory of mind agents with RFX-BMS has been previously used by us in [27] to accurately recover the level of theory of mind reasoning of Devaine’s Bayesian theory of mind agents [5]. This indicates that this method can overcome some of the biases in the designer’s choice of how to implement theory of mind. After all, the agents of one designer accurately modeled the theory of mind abilities of other, independently designed theory of mind agents.

5 Experimental Setup

To determine whether interacting with artificial theory of mind agents encourages the use of theory of mind, we let human participants play the Mod game against artificial theory of mind agents. Participants played the two-player Mod game with 24 actions, as described in Section 2.

5.1 Participants

Sixteen participants were included in this study, of which eight were male and eight were female, all students and all over the age of 18 ($M = 21.5$, $SD = 2.3$). The experiment was conducted in English, and all participants were sufficiently skilled in reading and understanding the English language, as they were all students of the University of Groningen, where an admission requirement is a sufficient proficiency in the English language.

Before starting with the experiment, all participants gave informed consent about partaking and about the use of the data obtained by the experiment for the purpose of this study.

5.2 Experimental design

Each participant played the Mod game against four different computer opponents: a ToM_1 agent, a ToM_2 agent, a ToM_3 agent, and an agent whose order was randomized each round. This randomizing agent would randomly select to respond as if it were a first-order, second-order, or third-order theory of mind agent in each round. That is, during a block of twenty rounds, the randomizing agent would randomly select a reasoning strategy twenty times. The ToM agents in the experiment did not exhibit any behavioral noise (i.e. $\beta \rightarrow 0^+$) and were set to learning speed $\lambda = 0.5$.

Note that participants never played against a ToM_0 agent. During a separate pilot study, it was discovered that ToM_0 agents and ToM_1 agents exhibit the same behavior when playing the Mod game against a human participant. The ToM_0 agent believes that the best predictor for a participant’s future behavior is the participant’s behavior in the most recent round. As a result, the ToM_0 agent tends to select the number that is 1 higher than the number last chosen by the participant.

The ToM_1 agent, on the other hand, believes that the opponent wants to win the game. By taking the perspective of the opponent, the ToM_1 agent believes that the opponent will choose the number that is 1 higher than his own last choice. For example, suppose that the agent chose 23 in the last round and the participant played 24. In this case, the ToM_1 agent believes that the opponent is going to play 24 again, since the ToM_1 agent believes that the participant is a ToM_0 agent who believes that the agent is going to play 23 again. Following this reasoning, the agent decides to play 1, which is exactly 1 higher than the participant’s previous choice (24).

Whenever the participant wins from the agent, the participant has chosen the number that is 1 higher than the number chosen by the agent. In this case, the behavior of a ToM_0 agent (choose the number that is 1 higher than the participant’s last choice) is the same as that of a ToM_1 agent (choose the number that is 2 higher than the agent’s own last choice). In both cases, the agent chooses to play the number that is 1 higher than the participant’s previous choice. When the participant wins consistently, the behavior of a ToM_1 agent is therefore almost indistinguishable from the behavior of a ToM_0 agent. Due to this effect, we decided not to include the ToM_0 agent in our experiment.

Each block consisted of twenty rounds of the Mod game, in which participants played against the same opponent for all twenty rounds. Between blocks, the opponent was changed to a ToM agent of a different order. The order in which participants faced the different opponents was randomly drawn from four possible sequences: $[?, 1, 2, 3]$; $[3, ?, 1, 2]$; $[2, 1, 3, ?]$; and $[1, ?, 3, 2]$, where the question mark (?) represents a randomizing agent, whose order of theory of mind reasoning was randomized each round. The different sequences were chosen to rule out the effect of sequence on the performance. Participants were informed about the ToM order of the opponent they were playing against, except in the blocks in which they faced the randomizing agent. During the rounds against the randomizing agent, the order of the opponent was not shown to the participants, in order to see if the participants’ behavior also changed if the ToM order of the opponent was not known. By inform-

ing the participants, we not only aimed to inform the participants that they were playing against an intelligent opponent, but also to entice them to think about their choices and strategy.

Each participant played two repetitions of four blocks, so that in total, a participant played forty rounds of the Mod game against each opponent. The sequence did not change within participants, so a participant faced a certain sequence of agents twice. A certain opponent was played against for twenty rounds before the agent's *ToM* order changed. The number of twenty rounds per opponent was chosen because people typically need many trials before showing higher-order reasoning behavior [10].

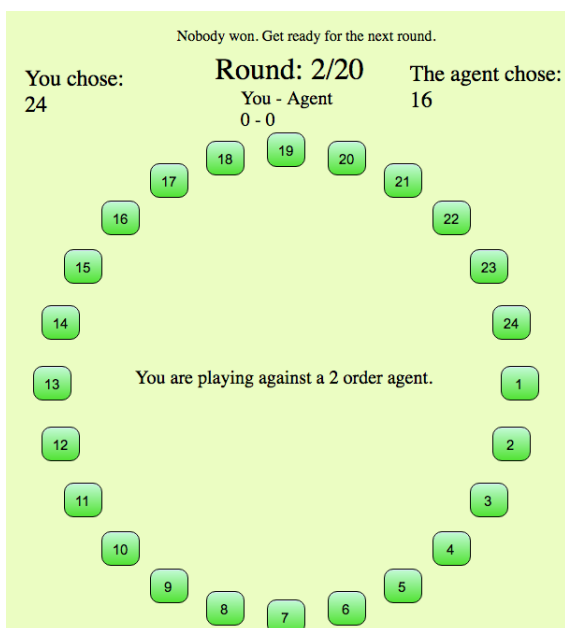


Fig. 2: Interface of the Mod24 Game experiment.

5.3 Procedure and materials

The experiment was run entirely on a MacBook. Since the experiment was web-based, only a browser (Google Chrome) was used. First, the participants read some short information about the theory of mind and the different orders that were used in this experiment (ToM_1 , ToM_2 , and ToM_3). Note that while this procedure may have primed participants to make use of theory of mind strategies, evidence from Marble Drop experiments shows that even in this case, participants may have difficulties implementing higher-order theory of mind strategies [11,17]. Participants then read an explanation of the experiment itself, including the rules of the Mod

game and an explanation of the interface. Before the experimental rounds started, the participants completed three test rounds to confirm they understood the interface. As the participants started with the experimental trials, they saw an interface with twenty-four buttons, numbered from 1 to 24, placed in a circle (see Figure 2). The placement of the buttons was constant throughout the whole experiment. The interface also showed what level of theory of mind the agent used (except during the randomizing agent blocks). The participants could also see how many rounds they had already played, how many rounds they would play against the same opponent, the current score of both players, and the chosen actions of both players in the previous round of play.

At the end of each block, participants were informed that they would continue playing against a new opponent. After four blocks, participants could take a break before continuing with the next four blocks. Once all eight blocks were finished, another pop-up was shown, informing the participants that the experiment was finished. Upon finishing the experiment, the participants were thanked for their cooperation and received payment. Each participant was equally compensated for their effort: the reward was not dependent on the points obtained during the experiment.

5.4 Data collection

During the experiment, the following variables were recorded. The reaction time (the time it took the participant to choose a number), the number chosen by the participant, and the number chosen by the agent in the current round. We also recorded the number chosen by the participant and the agent in the previous round; this was done to see whether or not there exists a relation between what the opponent chose previously and what the participant does next (and vice versa). The *ToM* order of the agent was also kept track of, as well as the number of wins for the participant, and the number that the participant chose that led to a win.

The data obtained was divided into groups per opponent and the differences in data between these groups were compared. Variations in rate (differences between the players' previous and current number) were compared to see if there was a correlation between the *ToM* order of the opponent and the rates the players used. These differences could indicate that participant behavior changed per opponent.

To investigate this further, an estimate was made about how likely it was that the participant data corresponded to certain pre-defined strategies. On these likelihoods, a random-effect Bayesian model selection (see Section 4) analysis was executed to determine which

part of the population used a certain strategy. The RFX-BMS was executed over the participant data of the whole experiment, as well as over the participant data per different *ToM* opponent. This was done to see whether the strategies that were used by the participants varied between the different opponents.

6 Experiment results

6.1 Agent behavior

Figure 3 shows the estimated strategies of each of the four agents in our experiment. Note that for the *ToM*₁, *ToM*₂, and *ToM*₃ agents, the RFX-BMS estimation correctly classifies agent behavior as consistent with the corresponding order of theory of mind reasoning. Moreover, the randomizing agent is classified as using a strategy that is approximately equally consistent with all strategies. This shows that the RFX-BMS estimation can accurately distinguish different order of theory of mind reasoning, and also considers the randomizing agent to be unpredictable.

In particular, Figure 3 shows that, although it is possible for a higher-order theory of mind agent to behave as if it were a lower-order theory of mind agent, such agents are not underestimated by RFX-BMS estimation. This means that for each agent, there are rounds in which the actions of the agent are very unlikely to be a result of reasoning at a lower order of theory of mind, which prevents underestimation of the agent. At the same time, the abilities of the randomizing agent are not overestimated as being consistent with third-order or fourth-order theory of mind.

Figure 4 shows the number of wins for each of the artificial *ToM* opponents, out of a possible 20. Note that in each round, exactly one out of 24 possible actions gives a score of 1. All other actions give a score of 0. As a result, chance level performance of a given round is $\frac{1}{24}$. Across 20 rounds, chance level performance is therefore $\frac{20}{24}$. As the figure shows, the *ToM*₁ agent performed especially poorly against the human participants, and obtained a median score of 0/20. The random order *ToM* agent performed better with a median score of 1/20. However, only the *ToM*₂ agent (median score of 3/20) and the *ToM*₃ agent (median score of 2/20) scored significantly higher than chance performance (sign test: $\hat{p} = 0.9375, p < 0.006$ and $\hat{p} = 1, p < 0.001$, respectively), against a human opponent.

6.2 Human behavior

In this section the results of the experiment are discussed, with a focus on how the participants reacted to the artificial *ToM* agents.

6.2.1 Overall results

In this section, the results of our experiment as a whole are discussed, aggregated over the different *ToM* orders of the opponent. As discussed in Section 5, participants played the Mod game against four agents with varying orders of *ToM* reasoning.

In order to test whether or not a certain sequence of opponent appearance was harder or easier than any of the other sequences, an ANOVA was executed over the participant scores per sequence. No significant influence of the sequence on the participant scores was found ($F(3, 60) = 0.598, p = 0.616$). That is, there is no reason to believe that the sequence in which the agents appeared affected the performance of the participants in any way. In the remainder, we therefore present results that are aggregated across the different sequences.

The overall behavior of the participants during the experiment is depicted in Figure 5. Figure 5a shows the frequencies of the numbers chosen by the participants (red line) and what the frequencies of the participants' choices would have been if they behaved randomly (green line). This figure shows that there was no clear preference for certain numbers, meaning that the participants behaved in an approximately random fashion. In addition, Figure 5b shows the rate, which is the difference between the participant's current and previous choice. For example, a rate of 3 means that the participant chose a number that was 3 higher than the number she previously chose. Figure 5b indicates that a rate of 2 is chosen the most over the course of the whole experiment. Overall, participant rates typically were between 0 and 5. All in all, this means that participants mostly chose a number that was two higher than their previous choice, and sometimes picked the same number they chose previously, or a number that was 1, 3, 4, or 5 higher than their previous choice.

6.2.2 Performance and reaction times

Figure 6 shows the performance of the participants per order of theory of mind of the opponent as the total number of wins. Since each participant played against the same opponent twice, each participant represents two data points in Figure 6.

The figure suggests that participants could relatively easily win from the *ToM*₁ opponent, while it was harder

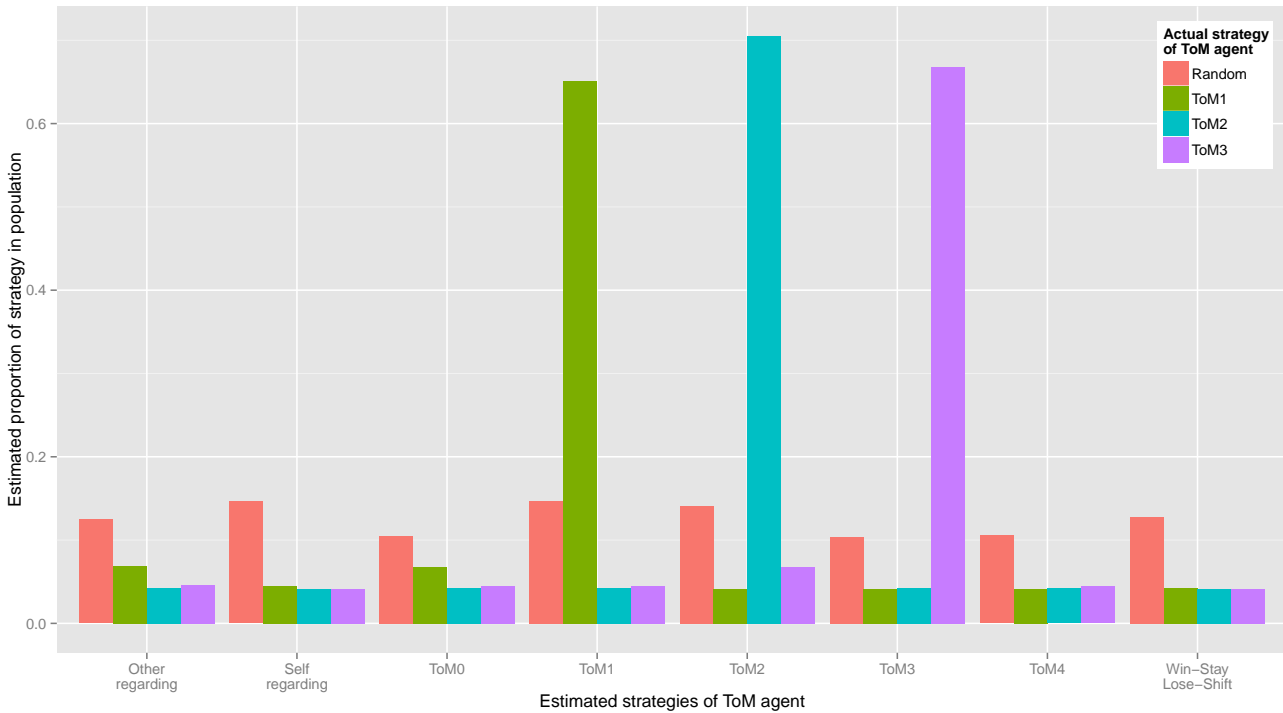


Fig. 3: Estimated strategies of the artificial *ToM* agents in the experiment.

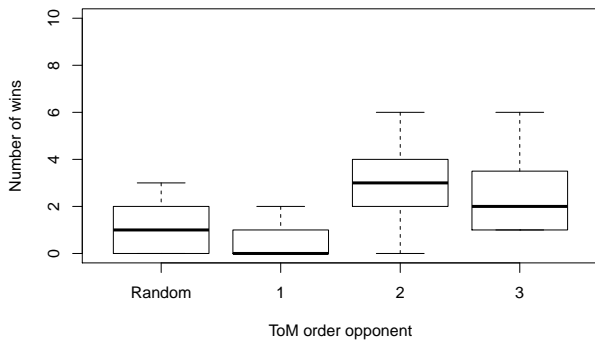


Fig. 4: Number of wins for each of the artificial *ToM* opponents.

to win from the higher-order theory of mind opponents, and almost impossible for the participants to win from the random order *ToM* opponent. A *ToM*₁ opponent leads to a relatively high number of wins, with a median participant score of 16/20. As the order of theory of mind of the opponent increases, the score of the participant decreases. Against a *ToM*₂ opponent, participants score a median 5 out of 20 points, while they only obtain 3 out of 20 points when facing a *ToM*₃ opponent. The lowest performance was observed when

the participants played against the random order *ToM* opponent, where participants achieved a median score of 1/20. In fact, when playing against the random order *ToM* agent, participants did not achieve scores that were significantly higher than chance level performance ($\hat{p} = 0.6875, p = 0.415$).

The influence of the *ToM* order of the opponent on the performance of the participants was found to be significant ($F(3, 60) = 66.34, p = 2.2e^{-16}$). Further testing showed that the success rate during the *ToM*₁ blocks was significantly higher than during any of the other blocks ($p = 2.2e^{-16}, p = 2.2e^{-16}$ and $p = 2.2e^{-16}$ for the *ToM*₂, *ToM*₃, and random order *ToM* blocks respectively). In addition, participants scored significantly higher when facing a *ToM*₂ agent or a *ToM*₃ agent than when they played against the randomizing agent ($p = 5.651e^{-12}$ and $p = 3.953e^{-13}$, respectively). However, participants' scores against the *ToM*₂ agent did not differ significantly from those against the *ToM*₃ agent ($p = 0.7258$).

Whether or not the reaction times of the participants differed per opponent was also investigated. It was hypothesized that participants that engaged in increasingly higher orders of theory of mind reasoning would also have increasingly longer the reaction times, e.g. reasoning in *ToM*₂ might take longer than reasoning in *ToM*₁. As the level of reasoning of the opponent

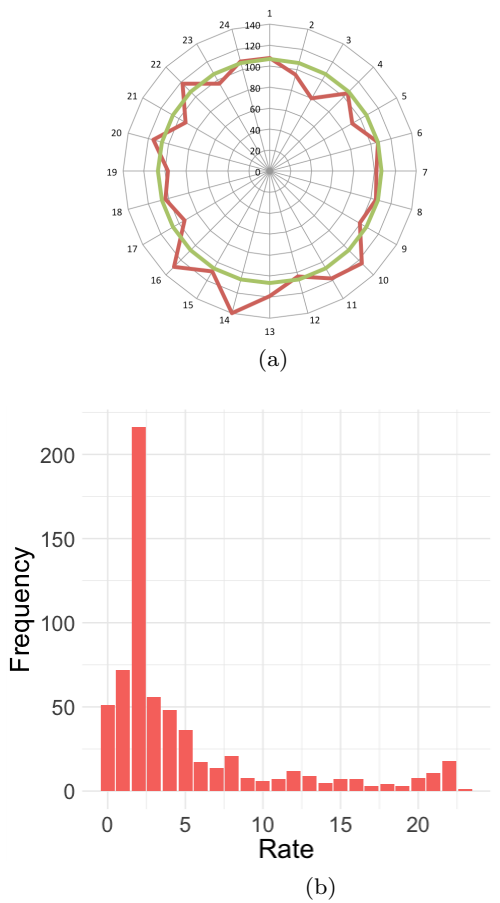


Fig. 5: Participant (a) choices and (b) rates during the whole experiment. The green line indicates the expected outcome of random behavior.

increases, ideally the level of reasoning of the participant increases as well, which may result in incrementing reaction times. To test this, an ANOVA was executed over the logarithmic reaction time data² of the participants, which are depicted in Figure 7.

With $p < 2e^{-16}$ and $F(3, 2556) = 34.85$ we found statistical evidence that the order of the opponent did influence the reaction times of the participants. This means that the reaction times of the participants can differ significantly per opponent. Post-hoc analysis was performed to see where the differences in reaction times per opponent lie. The analysis showed that every pair differed significantly, which indicates that no matter which opponents are compared, the reaction times of the participants differed in all conditions significantly from the other conditions.

The average reaction times of the participants when playing against a ToM_1 agent was 3693 ms, the ToM_1

² The logarithmic transform was used due to the skewed distribution of reaction times.

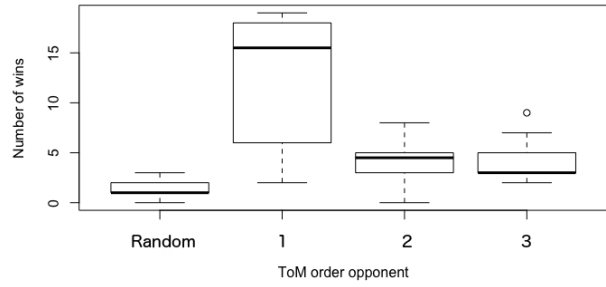


Fig. 6: Number of wins for the participant per ToM order of the opponent.

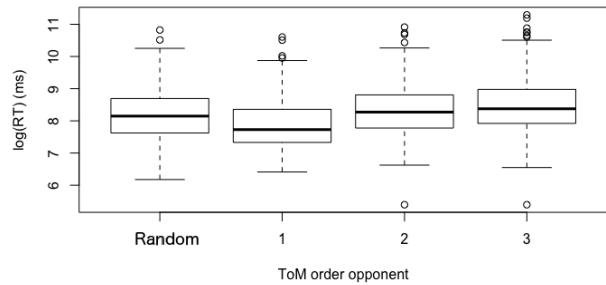


Fig. 7: Reaction times of the participant per ToM order of the opponent on a logarithmic scale.

agent was also the opponent that led to the quickest reactions. The overall reaction times during the trials against a second-order ToM opponent were longer than during the trials against a first-order ToM opponent with an average reaction time of 5653 ms vs. 3693 ms. During the ToM_2 opponent trials, more outliers were observed, and the maximum and minimum reaction times were more scattered than in the trials against a ToM_1 agent.

The reaction times when playing against a ToM_3 opponent were even more scattered and many outliers were observed during these trials. The average reaction time of the participants whilst playing against this opponent was 6684 ms, which is longer than the average reaction time during the ToM_1 trials, but seems close to the average of the trials against the ToM_2 opponent.

The average response time of the participants during the trials when the random order ToM agent was the opponent was 4836 ms, which is smaller than the reaction times during the ToM_2 and ToM_3 agent trials, but more than the response times during the ToM_1 opponent trials.

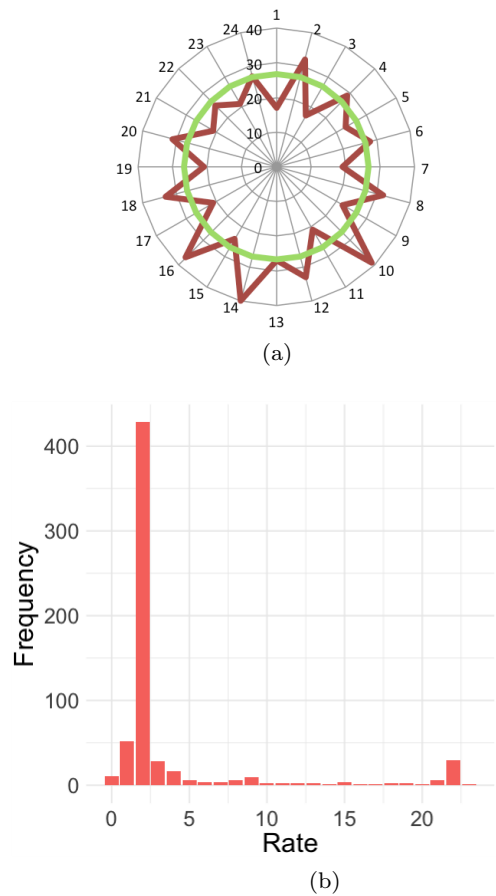


Fig. 8: Participant (a) choices and (b) rates during the ToM_1 opponent blocks. The green line indicates the expected outcome of random behavior.

6.2.3 ToM_1 opponent

In this section the results will be discussed that were obtained when the participants played against a ToM_1 opponent. The behavior of the participants can be seen in Figure 8. As can be seen in Figure 8a, participants favored even numbers over odd numbers when playing against a ToM_1 agent. They did not behave randomly, the red line (participant behavior) deviates from the green line (random behavior). Figure 8b shows the frequencies of different rates chosen by the participants. A rate of 2 was used the most when playing against a ToM_1 agent, this means that participants mostly chose a number that was 2 higher than their previous choice. The rate usage of the participants was investigated further, by looking at what rate each individual chose the most out of the trials played against a ToM_1 opponent. Every one of the 16 participants chose a rate of 2 most often during these trials. This finding seems to be in accordance with the aggregated frequency rates, observed in Figure 8b.

Compared to the behavior during the whole experiment (see Figure 5b), it can be seen that the rates when playing against a ToM_1 agent vary less. In the overall data, more spikes were observed, which is not observed here. Furthermore, the gradient of the rate figure when playing against a ToM_1 opponent also differs in comparison with the rate gradient of the whole experiment. There is also a larger difference in frequencies with regard to the preferred and ill-favored numbers.

According to the RFX-BMS (green bars in Figure 12), when playing against a ToM_1 agent, an estimated 25.8% of the participants made use of a ToM_2 strategy, while another estimated 22.4% of the participants used first-order theory of mind. It seems that the ToM_2 strategy was the best strategy (according to the participants). This also makes sense in relation to the theory, if you want to win a ToM game, it is most beneficial to think exactly one step further than your opponent does, so think in the second-order if the opponent thinks in the first-order. However, against a ToM_1 agent, a participant may also win using first-order theory of mind. Once the ToM_1 agent has lost all confidence in her first-order theory of mind, she will behave as if she were a ToM_0 agent. After this point, a participant using first-order theory of mind will win all future rounds.

Note that Figure 8b shows that participants often chose to play the number that was 2 higher than their previous choice, which suggests that participants may have used a self-regarding strategy with $k = 2$. However, the RFX-BMS results in Figure 12 show that participants were poorly described as using a self-regarding strategy. This is because the self-regarding strategy with $k = 2$ predicts that whenever a player fails to choose the number that was 2 higher than his previous choice, he will randomly choose one of the remaining numbers. The theory of mind strategies, on the other hand, predict that this player will choose a number that is slightly higher than either his own previously chosen number or is slightly higher than the previous choice of the opponent. Our RFX-BMS results suggest that the latter fits participant behavior better.

6.2.4 ToM_2 opponent

The participant behavior for the rounds against a ToM_2 agent can be seen in Figure 9. The red line in Figure 9a shows that the participants had no clear preference for certain numbers during these rounds. As is indicated by the small deviation of participant behavior (red line) from random behavior (green line), participant choices were approximately randomly distributed. Figure 9b shows the rate frequencies during the ToM_2 opponent rounds. During these rounds, a rate of 2 oc-

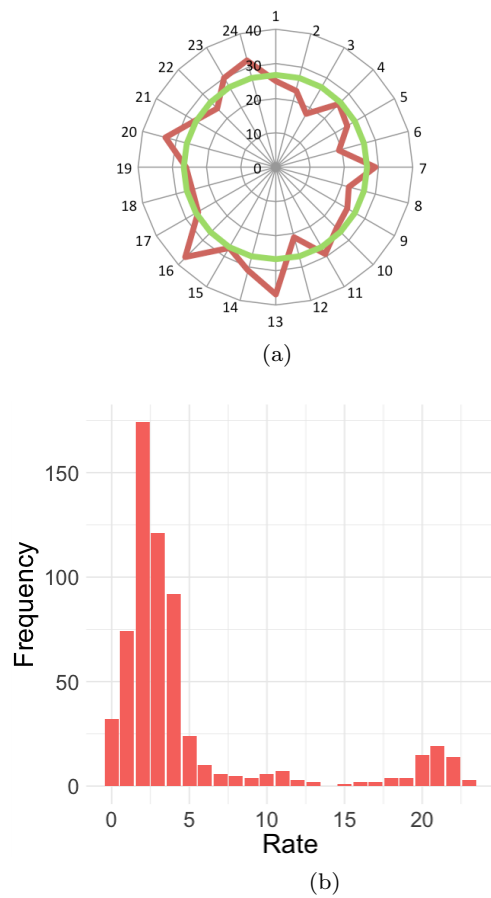


Fig. 9: Participant (a) choices and (b) rates during the ToM_2 opponent blocks. The green line indicates the expected outcome of random behavior.

curred the most, however, in Figure 9b it can also be seen that rates 3 and 4 occurred quite often as well. When the participants played against the ToM_1 agent, each of the 16 participants chose a rate of 2 most often, while during the trials against a ToM_2 agent, other rates than just 2 were chosen most often by some of the participants as well. When playing against a ToM_2 opponent, the most chosen rate per participant is still 2, however, some participants also chose a rate of 1, 3, or 4 the most.

When playing against a ToM_2 agent, participants mainly played according to first-order, second-order, or third-order theory of mind (blue bars in Figure 12). Herein a difference with the strategy for the ToM_1 opponent can be seen. The difference in ToM order of the opponent led to a different strategy that occurred the most in the population.

The results discussed in this section, combined with the results discussed in Section 6.2.1 indicate a difference in behavior of the participants per opponent, mainly in rate and in performance. In comparison with

the rounds played against a ToM_1 agent, we also see differences in strategies that explain the population the best. This indicates that the opponent against which participants are playing does influence their behavior and strategy.

6.2.5 ToM_3 opponent

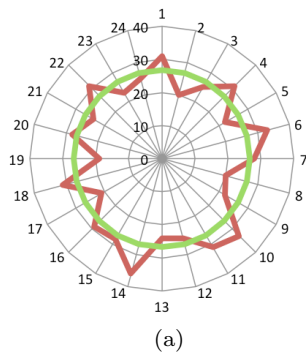
The participant behavior that was observed during the rounds played against a third-order ToM opponent can be found in Figure 10. The frequencies of the numbers chosen by the participants are shown in Figure 10a. This figure shows that the participant behavior (red line) again deviates from random behavior (green line), but overall the participants' behavior is approximately random.

The rate with which participants changed their action during the ToM_3 rounds can be seen in Figure 10b. The rates that are chosen the most frequent are the same as when the participants played against a ToM_2 agent, namely rates of 1, 2, 3, and 4. However, during the ToM_3 opponent rounds, it seemed that a rate of 4 was chosen more often than during the second-order ToM opponent rounds. These differences are more clear when looking at the individual preferences of the participants and what rate each participant chose the most. It was found that when playing against a ToM_3 opponent, a rate of 4 was chosen the most by the largest part of the population, whereas when playing against a ToM_2 agent, the majority of the participants still chose a rate of 2 the most. This indicates that while a rate of 2 occurred the most during these trials overall, when looking at the most chosen rate per participant a rate of 4 is chosen the most. This means that during the ToM_3 opponent trials, some participants might have had a strategy that entailed choosing a number that was 4 higher than their previous choice.

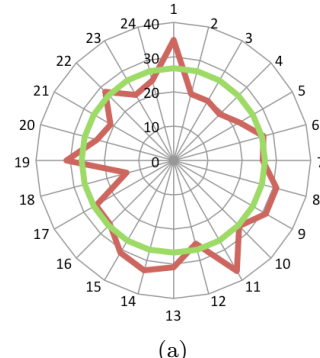
The strategy usage during these rounds was also investigated with a RFX-BMS analysis, of which the results are in the purple bars of Figure 12. This figure shows that the ToM_3 and ToM_4 strategies were the strategies that best explained the largest percentages of the population, 22.9% and 21.2% of the population respectively. Interestingly, those participants that were classified as using a higher order of theory of mind reasoning also obtained higher scores on average.

6.2.6 Random order ToM opponent

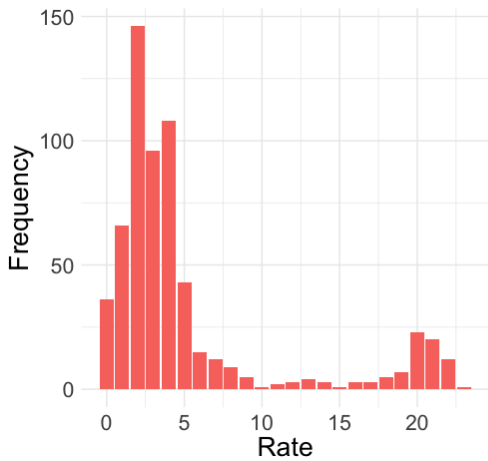
Figure 11 shows the behavior of the participants during the rounds where the ToM order of the opponent was randomly reassigned in each round. The frequencies of the numbers chosen by the participants can be seen



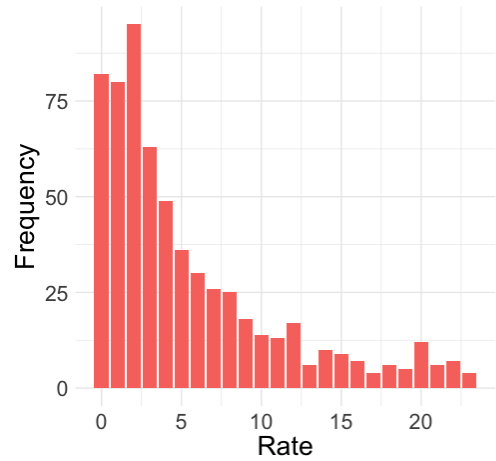
(a)



(a)



(b)



(b)

Fig. 10: Participant (a) choices and (b) rates during the ToM_3 opponent blocks. The green line indicates the expected outcome of random behavior.

Fig. 11: Participant (a) choices and (b) rates during the random order ToM opponent blocks. The green line indicates the expected outcome of random behavior.

in Figure 11a. This figure illustrates that the behavior of the participants during these rounds was approximately random. The rates displayed in Figure 11b show that the rates during the random order ToM opponent are distributed differently than the rates in the rounds against the other opponents. In contrast to the other blocks, during the random order ToM agent block, there seemed to be more of a preference to stay on the same number (a rate of 0). However, a rate of 2 is still the most frequent one, just as in the other blocks. This figure also shows that higher rates occurred more often during these trials as well (e.g. rates of 6, 7 or 8). This trend was also observed when looking at the individual data, where it was shown that some of the participants chose higher rates the most (e.g. rates of 5 or 7).

The results of the RFX-BMS in Figure 12 (red bars) show that among the ToM strategies, the ToM_2 strategy best describes participant behavior. However, when playing against the randomizing agent, participant behavior is better described as other-regarding or as a win-stay, lose-shift strategy.

Figure 12 also shows that, unlike when participants play against the ToM_0 , ToM_1 , or ToM_2 agent, participant behavior against the random ToM opponent is better explained by behavior-based strategies than it is by theory of mind strategies. While a sizable proportion of the population is still estimated to use second-order theory of mind, the other-regarding and win-stay-lost-shift strategies are estimated to account for more than 20% of the population each.

These results indicate that it was very hard for participants to decide on a strategy that led to many wins against this opponent. The performance during these trials was very low and the rates deviate from the rates during the trials in which the ToM order of the agent was fixed. Note, however, that the participants were not outsmarted by the random order ToM opponent. When playing against this opponent, both players obtained low scores.

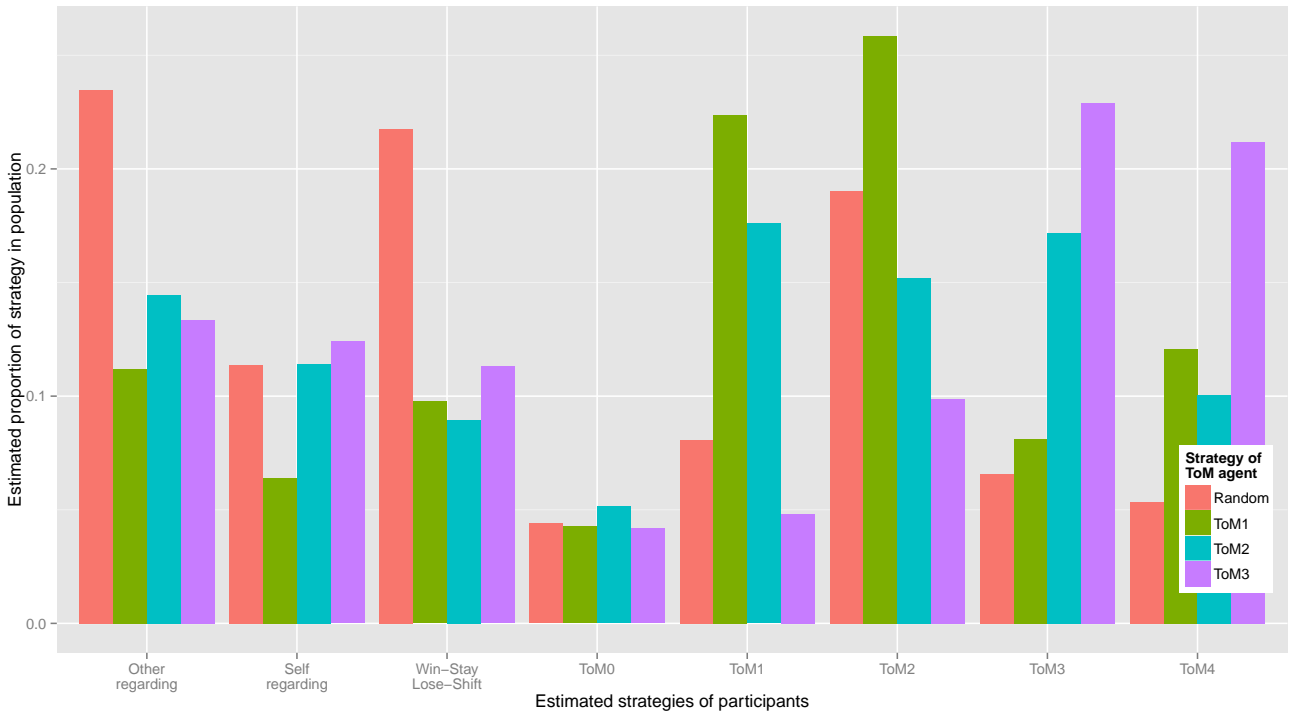


Fig. 12: Estimated strategy use of participants in the Mod game across the four different opponent types.

7 General discussion and conclusion

In this paper, we aim to determine (1) to what extent participants use theory of mind reasoning when playing against artificial agents, as well as (2) to what extent artificial agents can encourage the use of higher-order theory of mind by participants. We do so by letting participants play the Mod game [9,28] against artificial theory of mind agents.

In our experiment, a large proportion of participants playing the Mod game against theory of mind agents is best described as making use of higher orders of theory of mind. Participants that faced a first-order theory of mind agent relied mostly on first-order or second-order theory of mind themselves, while participants that played against a third-order theory of mind agent were better described as using third-order or fourth-order theory of mind. Moreover, these theory of mind strategies were found to describe participant behavior better than simpler behavior-based strategies, such as always choosing the number that is 2 higher than your previous action. However, when playing the Mod game against a randomizing agent, who randomly selected to play as if it were a first-order, second-order, or third-order at the start of each round, participants were better described as relying on such simpler behavior-based strategies.

In strategic games, participants are typically found to rely on low orders of theory of mind, and to be slow to adjust their level of theory of mind reasoning to more sophisticated opponents [11, 3, 32, 10]. Earlier empirical research suggests that the use of first-order and second-order theory of mind in games can be facilitated by creating a believable story or insightful visual representation around an abstract problem [5], by creating a clear competition or negotiation setting [10, 31], or by providing stepwise training from games that require zero-order *ToM* to second-order *ToM* games, as we did in [25].

Our results in the Mod game suggest that participants even make use of an unprecedented *fourth-order* theory of mind reasoning when playing against a higher-order theory of mind opponent in the Mod game, even though they only faced each opponent for twenty consecutive rounds of play. Moreover, participants that were classified as using higher orders of theory of mind tended to obtain higher scores. In future work, it would be interesting to determine to what extent participants exhibit the same behavior when facing more than one opponent at the same time (i.e., for $n > 2$).

Additionally, our results in the Mod game show higher levels of theory of mind reasoning than results of similar experiments with the matching pennies game [27, 5]. One possible explanation is that adults in the match-

ing pennies experiment did not have enough time to engage in higher-order theory of mind reasoning. In our Mod game experiment, participants took an average 3700 ms to respond in rounds where they faced the least sophisticated opponent. In contrast, participants in the matching pennies experiment were given 1300 ms to make a decision. Given this time pressure, participants may have decided to rely on simple behavior-based strategies.

Alternatively, although the Mod game is arguably more difficult than matching pennies, it is easier to distinguish strategies from one another in the Mod game than in matching pennies. In matching pennies, players only have two possible actions to choose from. As a result, the choice of a particular action gives little information about the underlying strategy. In our Mod game experiment, however, there are 24 possible actions, which allows players to more easily interpret the actions of their opponent and draw conclusions about the underlying strategy. In addition, participants were informed about the abilities of their artificial opponents, which may have helped them to identify observed behavior of the opponent.

The unexpectedly high level of theory of mind reasoning of participants can also be partially explained by the representation of the game. In our definition of theory of mind agents, zero-order theory of mind agents reason about opponent actions. However, this ignores the experimental setting that participants were confronted with, in which actions are meaningfully arranged in a circle. This may encourage people to think about the game in terms of the change of action (i.e. rate) rather than the specific choices that were made. A zero-order theory of mind agent that thinks in terms of rates rather than choices would exhibit behavior similar to first-order theory of mind agents that think in terms of choices. That is, the representation of the zero-order theory of mind model is important in determining at which order of theory of mind participants are reasoning [17].

Theory of mind plays a fundamental role in many social skills, and especially communication. Using a combination of language, prosody, body language, and gestures, the speaker attempts to find the best way to convey a certain meaning to the hearer. Meanwhile, the hearer tries to find the best interpretation for a given utterance. For efficient communication, an accurate estimation of the other's level of theory of mind reasoning is vital. In this paper, we have shown that the use of artificial theory of mind agents provides a modality-independent way of obtaining such an estimate. In future research, it would be interesting to see whether robots using multi-modal communication are

more effective than software agents at enticing people to employ higher orders of theory of mind.

Conclusion

In our experiment, participants that played the Mod game against virtual agents capable of higher orders of theory of mind reasoning were estimated to engage in higher-order theory of mind reasoning themselves as well. This suggests that artificial agents can indeed encourage people to make use of higher-order social cognition and allow them to achieve better results. It would be valuable to extend this training approach using virtual theory of mind agents to other settings, such as cooperative games and coordination situations, in which being able to apply higher orders of theory of mind would also be beneficial for people.

Some existent virtual agents train people in social skills for which both first- and second-order theory of mind would be very important, such as training potential victims of doorstep scams to assess whether a scam is being attempted [16] and training police agents to avoid false confessions from suspects in an interrogation [2]. For such virtual training agents, it would be very useful to integrate our artificial agents' capabilities. Then the virtual training agents can both assess participants' levels of theory of mind and train them in using second-order theory of mind in adversarial situations. This will enable participants to apply useful complex reasoning, as in: "What does the person on the doorstep *intend* me to *believe*?" and "Could I have accidentally communicated something to the suspect that he should not *know that I know*?"

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. Adibsereshki, N., Nesayan, A., Gandomani, R., Karimlou, M.: The effectiveness of theory of mind training on the social skills of children with high functioning autism spectrum disorders. *Iranian Journal of Child Neurology* 9(3), 40–49 (2015)
2. Bruijnes, M.: *Believable Suspect Agents: Response and Interpersonal Style Selection for an Artificial Suspect*. Ph.D. thesis, University of Twente (2016)

3. Camerer, C., Ho, T., Chong, J.: A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119(3), 861–898 (2004)
4. Chevallier, C., Noveck, I., Happé, F., Wilson, D.: What’s in a voice? Prosody as a test case for the theory of mind account of autism. *Neuropsychologia* 49(3), 507–517 (2011)
5. Devaine, M., Hollard, G., Daunizeau, J.: The social Bayesian brain: Does mentalizing make a difference when we learn? *PLoS Computational Biology* 10(12), e1003992 (2014)
6. Devaine, M., Hollard, G., Daunizeau, J.: Theory of mind: Did evolution fool us? *PloS ONE* 9(2), e87619 (2014)
7. Franke, M., Galeazzi, P.: On the evolution of choice principles. In: *Proceedings of the Second Workshop Reasoning About Other Minds: Logical and Cognitive Perspectives*. CEUR Workshop Proceedings, vol. 1208, pp. 11–15 (2014)
8. Frey, S.: Complex collective dynamics in human higher-level reasoning: A study over multiple methods. Ph.D. thesis, Indiana University (2013)
9. Frey, S., Goldstone, R.L.: Cyclic game dynamics driven by iterated reasoning. *PloS ONE* 8(2), e56416 (2013)
10. Goodie, A.S., Doshi, P., Young, D.L.: Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making* 25(1), 95–108 (2012)
11. Hedden, T., Zhang, J.: What do you think I think you think? Strategic reasoning in matrix games. *Cognition* 85(1), 1–36 (2002)
12. 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)
13. Imuta, K., Henry, J.D., Slaughter, V., Selcuk, B., Ruffman, T.: Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology* 52(8), 1192–1205 (2016)
14. Kinderman, P., Dunbar, R.I., Bentall, R.P.: Theory-of-mind deficits and causal attributions. *British Journal of Psychology* 89(2), 191–204 (1998)
15. Liddle, B., Nettle, D.: Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology* 4(3-4), 231–244 (2006)
16. van der Lubbe, L., Bosse, T., Gerritsen, C.: Design of an agent-based learning environment for high-risk doorstep scam victims. In: *International Conference on Practical Applications of Agents and Multi-Agent Systems*. pp. 335–347 (2018)
17. Meijering, B., Taatgen, N.A., van Rijn, H., Verbrugge, R.: Modeling inference of mental states: As simple as possible, as complex as necessary. *Interaction Studies* 15(3), 455–477 (2014)
18. Mol, L., Krahmer, E., Maes, A., Swerts, M.: The communicative import of gestures: Evidence from a comparative analysis of human–human and human–machine interactions. *Gesture* 9(1), 97–126 (2009)
19. Mol, L., Krahmer, E., Maes, A., Swerts, M.: Adaptation in gesture: Converging hands or converging minds? *Journal of Memory and Language* 66(1), 249–264 (2012)
20. Premack, D., Woodruff, G.: Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1(04), 515–526 (1978)
21. Rappoport, A., Budescu, D.: Randomization in individual choice behavior. *Psychological Review* 104(3), 603 (1997)
22. Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J.: Bayesian model selection for group studies. *Neuroimage* 46(4), 1004–1017 (2009)
23. Stiller, J., Dunbar, R.I.: Perspective-taking and memory capacity predict social network size. *Social Networks* 29(1), 93–104 (2007)
24. Veltman, K., de Weerd, H., Verbrugge, R.: Socially smart software agents entice people to use higher-order theory of mind in the Mod game. In: *BNAIC 2017 Preproceedings*. pp. 253–267 (2017)
25. Verbrugge, R., Meijering, B., Wierda, S., van Rijn, H., Taatgen, N.: Stepwise training supports strategic second-order theory of mind in turn-taking games. *Judgment and Decision Making* 13(1), 79–98 (2018)
26. Wagenaar, W.: Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin* 77(1), 65–72 (1972)
27. de Weerd, H., Diepgrond, D., Verbrugge, R.: Estimating the use of higher-order theory of mind using computational agents. *The B.E. Journal of Theoretical Economics* 18(2) (2018)
28. de Weerd, H., Verbrugge, R., Verheij, B.: Theory of mind in the Mod game: An agent-based model of strategic reasoning. In: *Proceedings ECSI*. pp. 128–136 (2014)
29. de Weerd, H., Verheij, B.: The advantage of higher-order theory of mind in the game of limited bidding. In: Eijck, J.v., Verbrugge, R. (eds.) *Proceedings of the Workshop on Reasoning About Other Minds: Logical and Cognitive Perspectives*. pp. 149–164. CEUR Workshop Proceedings (2011)
30. 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)
31. de Weerd, H., Verbrugge, R., Verheij, B.: Negotiating with other minds: The role of recursive theory of mind in negotiation with incomplete information. *Autonomous Agents and Multi-Agent Systems*, 31, 250–287 (2017)
32. 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)