

# Thinking while playing, Theory of Mind in the stag hunt

(Bachelorproject)

Kenneth Muller, s2569434, s.k.muller@student.rug.nl,  
Harmen de Weerd\*, Trudy Buwalda\* and Burcu Arslan\*

July 14, 2016

## Abstract

The ability to understand that other people have different mental states, such as desires, beliefs, knowledge and intentions, which can be dissimilar to ones own, is called theory of mind. This study intends to find out if people make use of this ability in decision making while playing games, or if they use a different strategy. The game that was chosen is the stag-hunt game, in which a player need to cooperate with an AI opponent to gain the maximum amount of points. We set up an experiment in which fifteen participants each played 320 rounds of the stag hunt game, where the strategy of their opponent changed every ten rounds. We discovered that the participants do not show significant signs of the use of theory mind, in fact, theory of mind was used least frequently by all of the participants. However, a couple of strategies that the opponents in the game used, were abundantly used by the participants as well. From these results we can conclude that participants are less likely to use theory of mind in cooperative games, as opposed to using strategies that only consider the current game state and not the belief state of their opponent.

## 1 Introduction

### 1.1 Theory of Mind

When you play games against other people, it might be advantageous to wonder what your opponents might be thinking about. It is known that people are capable of reasoning about what other

people know and how they would respond with their given information. This is referred to as Theory of Mind (Premack & Woodruff, 1978).

To illustrate how Theory of Mind works, look at this example: Two people, Ann and Bob, play a game of rock-paper-scissors. Ann knows that most people generally start with rock, therefore Ann would think that playing paper would be the best option. In this case, Ann is not thinking about the thoughts of Bob, she is reasoning about a fact and not about any belief states, this is known as 0 order Theory of Mind. Bob also knows that people generally start with rock, and thinks that Ann knows this fact as well. If Ann would then make a decision based on this fact, she would play paper. With this in mind, Bob decides that going with scissors will help him win the game, Bob is applying first-order Theory of Mind here, since he is reasoning about Ann's belief state of a fact.

If Ann wanted to win this game of rock-paper-scissors, she should have reasoned about Bob's belief state. If Ann considered Bob's belief state, which is that Bob knew that Ann knew that people generally start with rock, she could have reasoned that Bob expected paper and would therefore play scissors. With this information, Ann could have played rock and won the game. In this case, Ann would be reasoning about what Bob believed to be Ann's belief state, this is one step ahead of the previous way of reasoning and is known as second-order Theory of Mind. This process could be repeated infinitely to create many orders of Theory of Mind.

It has been shown that humans are capable of using Theory of Mind while playing strategic games

---

\*University of Groningen, Department of Artificial Intelligence

and that within these strategic games people will generally attempt to employ an order of Theory of Mind that is one order higher than what they expect the opponent to use (Hedden & Zhang, 2002). Furthermore, it is also not too difficult to apply higher orders of Theory of Mind in strategic games so long as people are slowly introduced to it with proper instructions and training (Meijering et al., 2011).

## 1.2 Research question and hypothesis

Research has already shown that priming or training people on Theory of Mind will also allow them to easily apply higher orders of Theory of Mind (Meijering et al., 2011). The aim of this study is to find out if people will also attempt to reason through Theory of Mind if they are not primed or instructed to use it at all, or if they perhaps will apply different strategies. In the paper by Hedden and Zhang, participants played strategic games and they used Theory of Mind to reason about their opponents.

However, we will be using a different game than the games that were used by Hedden and Zhang, our participants will be playing a variant of the stag hunt. In this game they will be playing multiple rounds against an AI opponent which is not using Theory of Mind strategies. Instead, the opponents will change their strategies after every ten rounds. Will the participants still use Theory of Mind when they are not specifically told about Theory of Mind or reasoning whatsoever? If so, will they also be able to reason about the changes in strategy of their opponents? We hypothesize that if people play against opponents that apply clear strategies, people will still be capable of reasoning about the belief state of their opponents, meaning that people will still use Theory of Mind even without training.

Since we know that introducing people to Theory of Mind or training them to use Theory of Mind will allow them to properly apply this reasoning strategy in games, we will be careful to try and avoid this as much as possible before we let our participants participate. It was shown that people can apply several orders of Theory of Mind in strategic games, so we will let participants play the stag hunt, which also classifies as a strategic game.

What was also shown, is that people can easily apply higher orders of Theory of Mind when they are trained (Meijering et al., 2011). When we give the participants instructions on the experiment, we instruct them on how to play the game, and in which ways they can score points. They are also urged to score as many points as possible. They are not told anything about strategies or belief states of their opponents. This way, if people frequently use Theory of Mind during our experiment, we should be able to conclude that people can use Theory of Mind even without any introductions or training.

The question we will be asking while performing this experiment is: Will people use Theory of Mind in games, even without training or instructions on this method? We hypothesize that people will mostly apply Theory of Mind to reason about the belief state of the opponent and use this information to make a decision within the game. However, since we let our participants play with AI co-players that apply different strategies, we expect that they will also be quite likely to apply the same strategy as their opponent. This means that for all the trials where the participants play against an AI with strategy A, the likelihood that the participants use strategy A will also increase, whereas if they play against an AI with strategy B, they will be more likely to apply strategy B. The reasoning behind this is that to score optimally in the stag hunt, both hunters need to perform the same action. If the opponent is catching a rabbit, trying to catch a stag will net no points. On the other hand, if the opponent is attempting to catch a stag, the other player will get points for catching a rabbit, but cooperating to get the stag will result in more points.

## 2 Methods

### 2.1 Participants

While searching for participants, we mention that we are conducting an experiment on decision making in strategy games, avoiding the phrase Theory of Mind. For this experiment we had 15 participants, eight men and seven women, most of which were university students. All of the participants filled a consent form and they were compensated 4 euro's for participating.

## 2.2 The stag hunt

The game we intend to let people play in order to find out whether or not they will use Theory of Mind is a variation on the stag hunt. The stag hunt is a two-player game where two hunters have the choice of either catching a rabbit, which can be done alone, or catching a stag, which requires the hunters to cooperate. If a hunter wants to catch the stag which is worth 20 points, he requires the cooperation of the other hunter if he wants to succeed. On the other hand, the hunter can work alone and catch a rabbit, but this is worth only 10 points, which is half the amount of points a stag would give. This leads to a pay-off matrix which can be seen in Table 1.

As can be seen here, catching the stag is the most lucrative option, but this only works when both hunters attempt to catch the stag, otherwise attempting to catch the stag will result in 0 points. When either a rabbit or a stag has been caught, the round ends and a new round starts where both hunters can attempt to catch an animal again. Since the hunters will be playing multiple rounds of this game, it is important to be reasoning about which animal the other hunter will likely be trying to catch to succeed.

	Stag	Rabbit
Stag	20, 20	0, 10
Rabbit	10, 0	10, 10

Table 1: stag-hunt payoff matrix

We implemented a Java-based experiment of the stag hunt based on the model by Yoshida, Dolan, and Friston (2008). All of our participants played 32 blocks of 10 rounds each, resulting in a total of 320 rounds per participant. An example of the game can be seen in Figure 1.

We recorded the starting state of each round, which consists of the starting positions of the participant, AI and stag in the game, and the actions that a player takes during each round. The starting state always shows the player on the bottom row, the stag in the middle, and the AI opponent on the top row. Furthermore, there is a stationary rabbit on both sides of the middle row.

At the start of the game the player and AI are positioned in such a way that the distance from

both players is equal to the stag while both players are also equally close to a different rabbit. This results in a total of 5 different starting states. When it is the player’s turn to move, they can move in a direction (up, down, left, right) or they can decide to pass. Each move that a player takes during a round is seen as an action. So if a player plays a round and goes up two tiles, passes for one turn, and then moves left two tiles, after which it captures the rabbit, this will be indicated as ”up, up, pass, left, left, player catches rabbit”.

## 2.3 Procedure

Our participants were instructed to play the stag hunt game. One of the possible starting states of the game can be seen in Figure 1. Each participant started a round in a similar starting state to Figure 1. Participants always had a maximum of fifteen moves to finish a round and their remaining moves for a round were always shown on the top left side of the screen.

During the game, the order in which the characters move always remained the same: The stag made the first move, followed by the participant and then the AI. This order is repeated until someone caught a rabbit, when the maximum amount of moves had been reached, or when the participant and the AI caught the stag together. The stag is caught when both hunters are standing next to the stag and it is no longer able to move in any direction.

For each turn, the participant was allowed to either move in a direction or stand still. However, both options exhausted a move. The hunter AI is capable of using three different strategies, named cooperative, competitive and opportunistic. When the AI uses the cooperative strategy, it will always attempt to decrease the distance towards the stag. For the competitive strategy, the AI will instead move towards the nearest rabbit and attempt to catch it. The opportunistic strategy will attempt to catch the nearest rabbit unless both hunters are within three tiles of the stag. After every 10 rounds, the strategy of the hunter AI had changed. The participants were not told about this, instead they were expected to realize this and alter their strategy accordingly. After 320 rounds, the participants were notified that the game is done.

## 2.4 Design

In this experiment, we recorded the step that a participant took from each state in the game that the participant was in. A state is defined as the positions of the participant, the stag, and the opponent. For every state that the participant is in, we recorded the move that the participant made from that given state. This information was used to deduce the strategy that a player was employing from every state. We also measured if participants used different strategies in similar states if they were playing against an opponent that used a different strategy. So not only the state itself was important, but also the type of opponent that a participant played against.

## 2.5 Analysis

Analysis was performed with a method called RFX-BMS introduced by Rigoux, Stephan, Friston, and Daunizeau (2014). RFX-BMS stands for random effect Bayesian model selection, meaning that all models are treated as random effects. These effects can all differ between our participants, with an unknown population distribution. In our case, the models are the strategies that we expected our participants to employ. Before performing the RFX-BMS we defined the nine models that we wished to use for our analysis.

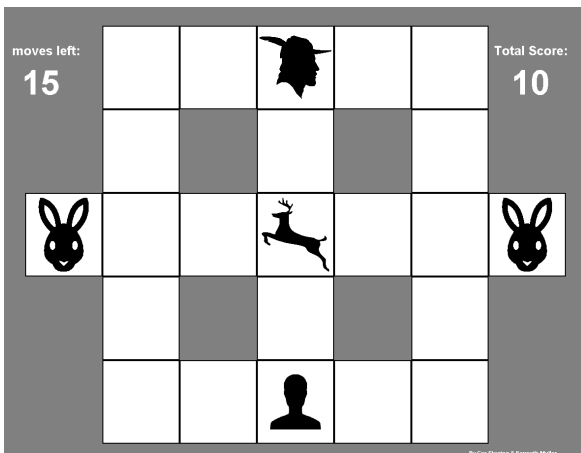


Figure 1: One of the five possible starting states of the stag hunt.

### 2.5.1 Theory of Mind models

First of all three of our models were Theory of Mind (ToM) models (de Weerd et al., 2013), ranging from zero-order ToM (ToM0) to second-order ToM (ToM2).

1) The ToM0 model does not reason about the beliefs of the opponent. Instead it believes that the game state that results from the previous one is purely dependent on what move the model itself makes. The model will be more likely to perform a move from a given state if it has made this move before and it lead to a winning outcome. On the other hand, a move will be less likely made again if this move lead to a loss.

2) The first-order ToM (ToM1) model reasons about the belief state of its opponent. When a ToM1 model reasons about a move, it will consider how the opponent will move from the resulting state as if the opponent were playing according to a ToM0 model. What differentiates the ToM1 model from the ToM0 model is that the ToM1 model considers whether or not the opponent is willing to cooperate. The ToM1 model attempts to reason about the belief state of the opponent, and if the model assumes that the opponent is only decreasing its distance to a rabbit, then the ToM1 model will not attempt to catch a stag, since the opponent is not trying to cooperate. While the ToM0 model would attempt to maximize its own score by determining which of its own options is optimal, the ToM1 model will also attempt to reason whether the opponent is willing to cooperate in order to maximize the score.

3) The second-order ToM (ToM2) model also reasons about the belief state of its opponent, but instead of considering the opponent as a ToM0 model, it will make a move while assuming that the opponent is playing according to a ToM1 model. The ToM2 model therefore assumes that the opponent can reason about the belief state of the model. This model then believes that if the model itself shows that it is willing to cooperate, the opponent will also be more likely to start cooperating, since the opponent reasons that the model is trying to catch the stag to maximize points as well.

### 2.5.2 Additional models

In addition to the ToM models, we used three different models, each representing different strategies: Random, Win-stay-lose-shift and Sticky.

4) The Random model has an equal likelihood to make every possible move from every given state, as the name already states it has a random chance of doing anything. This model never updates a belief state, so at every point in the game it keeps the same random chance of making any move. If this strategy is prevalent within the population it would mean that participants are not thinking about anything while playing the game, and are simply pressing buttons without reason.

5) The Win-stay, lose-shift model adapts its belief of a state depending on the outcome. If the model played a certain move in a state, and that eventually resulted in capturing the stag, the model is more likely to perform that move again. However, if that move lead to the AI capturing the rabbit, the model will be less likely to perform that move again. If this strategy has a high frequency it means that the participants are more likely to adopt a sort of trial and error strategy, trying moves until they seem to work.

6) The Sticky model is more likely to perform a move in a state if the player had performed this move before, regardless of the outcome. If this strategy is prevalent it means that the participants prefer making moves that they had performed before.

### 2.5.3 AI models

Finally, we also incorporated the three strategies that the hunter AI's used into the RFX-BMS as models to see if players would copy the strategies of their opponents: competitive, cooperative and opportunistic.

7) The competitive model will always make a move which takes it as close as possible to a rabbit. If there are multiple moves that decrease the distance, and these resulting positions all have an equal distance to a rabbit, then a random move out of that set is chosen. For example, if we look at the state in figure 1, every move except for standing still will cause the hunter to have a distance of three tiles from a rabbit, so any of the moves can be chosen.

8) The cooperative model will always make a

move which minimizes the distance between the hunter and the stag. Similarly to the competitive model, if there are multiple moves which lead to an equal distance from the stag, one of these moves will be chosen at random.

9) The opportunistic model is a combination of the previous two models, since it will almost always reduce their distance to the rabbit, unless both players are within three squares from the stag, in which case it will reduce their distance to the stag.

With all of these models, we performed the RFX-BMS on the data to determine which strategies are most prevalent within our population.

## 3 Results

We have not excluded any participants or trials from the analysis, which means that we had a total of 15 participants who each performed 320 trials. This resulted in a total of 4800 trials among all our participants for our analysis.

After performing the RFX-BMS on this data, all models were given relative frequencies. By adding all of these values and then dividing each of the separate values by the given sum, we obtain the likelihood that every strategy is used in percentages. These relative frequencies are depicted in Figure 2, where all models are shown on the x-axis, with the population frequency that uses a given model in percentages on the y-axis. Furthermore, we have added three extra figures, which present relative frequencies for the models from all trials where participants played against a specific AI strategy.

Our results in Figure 2 show that opportunistic is the most used strategy among our population. Furthermore, we see that the theory of mind models are the least used strategies among our participants. What we can tell from Figures 3, 4 and 5 is that regardless of the strategy that participants play against, they will always be most likely to use the opportunistic strategy. However, in Figure 3 we do notice that the relative frequency of competitive is equally high as that of opportunistic, the most prevalent strategy. The reasoning behind this is that against a competitive opponent, the hunter AI will always move towards the rabbit. The moment the participant figures this out, they should only attempt to catch a rabbit themselves. This should lead to the relative increase in frequency for

competitive which we see.

In Figure 4 we also see a slight increase in relative frequency for cooperative, while some other strategies drop in frequency, but the relative frequencies for competitive and opportunistic are still higher in comparison. Finally, in Figure 5 we see that the relative frequency of opportunistic strategy is much higher than the other strategies. Even though the relative frequencies of the opportunistic and competitive strategies are fairly high against an AI applying the same strategy, this is more likely to be a coincidence, since the relative frequencies of these two strategies are always the highest within the population.

## 4 Discussion

### 4.1 Hypothesis and theoretical implications

Our results suggest that instead of using theory of mind while playing strategy games, people are more likely to apply different strategies. These results seem to contradict our hypothesis concerning theory of mind, since we hypothesized that people would use theory of mind for decision making in strategy games, while our results showed that people preferred all other possible strategies over theory of mind.

What we can conclude from these results is that since people are most likely to apply the opportunistic strategy, they are more likely to reason about game states as opposed to the belief state of their opponent. To reiterate, if someone applies the opportunistic strategy, they will always move towards the rabbit unless both hunters are within three squares of the stag. If this is the most used strategy, then the participants are not reasoning about the belief state of the opponent, since the belief state of the opponent is not used for the opportunistic strategy, only the position of the opponent.

### 4.2 Literature, limitations and further research

Previous research had shown that people are capable of applying Theory of Mind in strategy games (Hedden & Zhang, 2002). However, our experiment

has shown different results, since zero until second order Theory of Mind were the least likely used strategies within our population. Furthermore, it was suggested that people had no difficulties learning to use Theory of Mind with proper instruction and training (Meijering et al., 2002). Perhaps this is also related to our findings, since we have shown that without proper instructions and training, people are less likely to use any form of Theory of Mind.

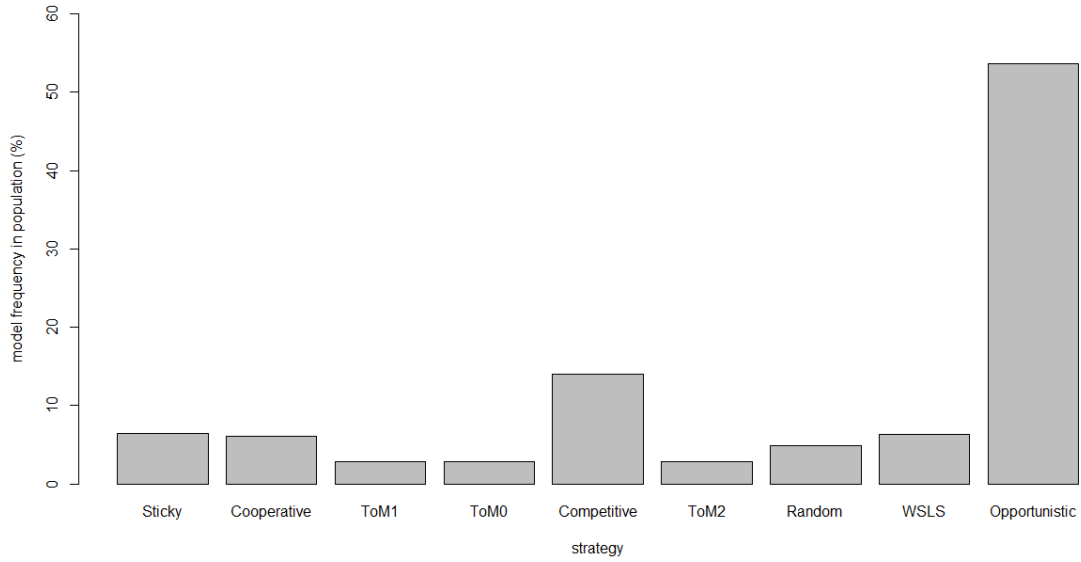
There are a couple of aspects that have potentially impacted our results. First of all, it is important to note that when we perform the RFX-BMS with the nine models mentioned in Section 2.5, we assume that those are the only nine strategies that the participants could have used. If we wanted to test if participants used a strategy or a higher order of Theory of Mind that is different from the ones that were mentioned we would have to implement this in the RFX-BMS as well. So there is a possibility that another strategy was used, which we had not considered.

Secondly, multiple participants mentioned that they would generally base their strategy around the first one or two moves of their opponent. This already means that they are not constantly reasoning about the belief state of their opponent, which rules out Theory of Mind for most game states. It is perhaps possible that participants used theory of mind at the beginning of the game, but used a different strategy for reasoning after the first move. This, once again, also strengthens our belief that people are more likely to reason about game states instead of the opponents' belief state.

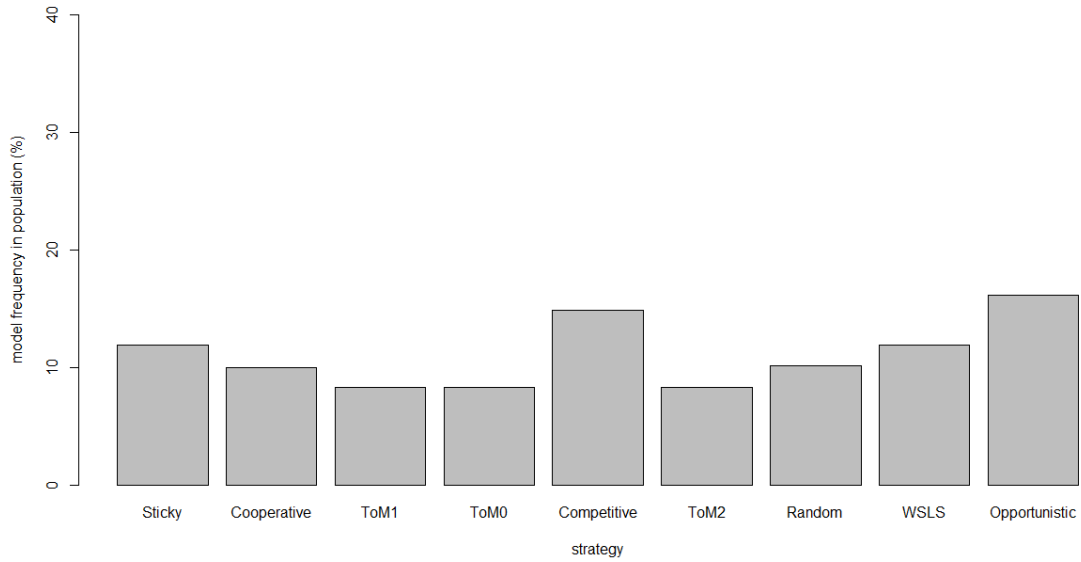
Finally, some participants expressed that they got fatigued while playing the game. All participants had to play the same game for 320 trials, lasting for approximately 30 minutes. This could lead to participants being less concentrated over time. This could lead to participants not altering their strategy any more over time, causing them to stick to something that worked for the first couple of trials.

Our results show that in our game, people are not likely to use Theory of Mind. However, this does not imply that people will not use Theory of Mind in any cooperative game.

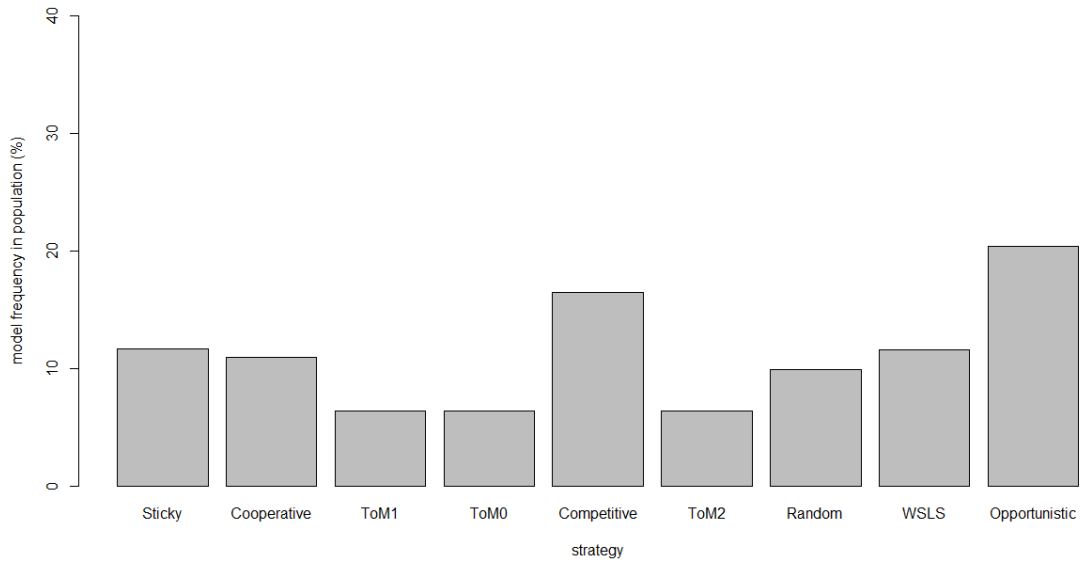
In our specific situation, the participants did not only need to reason about the belief state of their opponent, they also needed to bear in mind the location of the stag relative to the hunters and where



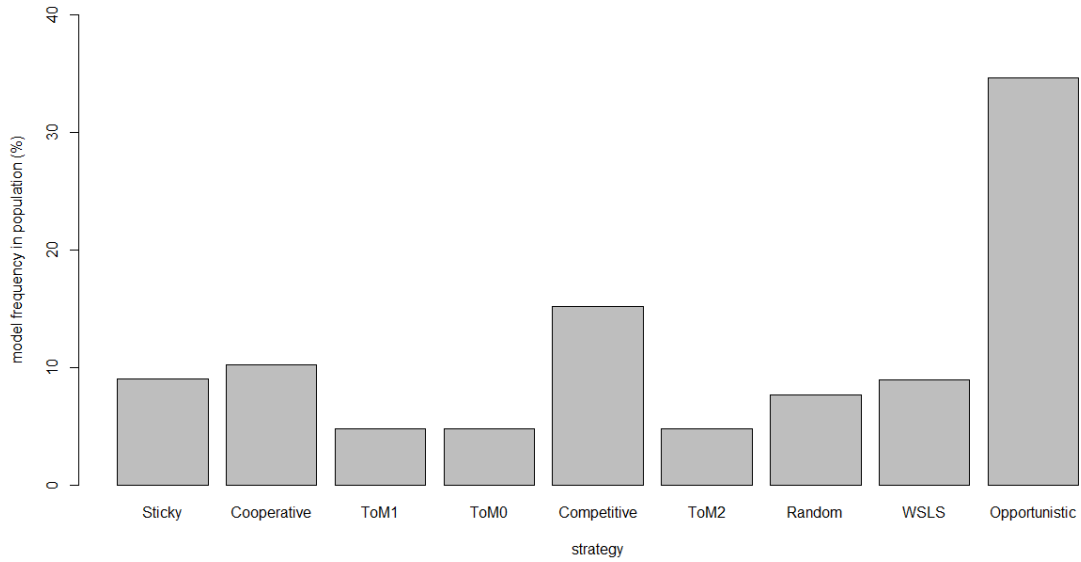
**Figure 2:** relative frequencies of all given strategies for all trials.



**Figure 3:** relative frequencies of all given strategies, only against competitive AI.



**Figure 4: relative frequencies of all given strategies, only against cooperative AI.**



**Figure 5: relative frequencies of all given strategies, only against opportunistic AI.**



it would move, in accordance with the amount of moves that the participant still had remaining. Perhaps in a game where there would be fewer variables, such that the participant can reason mostly about the belief state of their opponent, the chance of using Theory of Mind would improve.

Furthermore, in our situation, the AI could alternate between three preset strategies. The AI would swap between one of these strategies every ten trials, causing the participants to be unsure of how their opponents would react. Perhaps this caused people to not use Theory of Mind, since the belief state of the opponent was constantly changing. If we would let the participants play against any strategy for more than ten trials, they would perhaps be more likely to reason about the belief state of the opponent to apply theory of mind. Another possibility would be to let participants play against an AI which applies Theory of Mind itself. In this situation it would be optimal to apply an order of Theory of Mind which is higher than that of the opponent. In this case we would be able to tell if people will use a higher order Theory of Mind, if they would apply Theory of Mind at all.

To conclude, our results showed us that people are not likely to apply Theory of Mind while playing the stag hunt game. However, within a different game with fewer variables, or against different opponents in the same game, people may be more likely to use Theory of Mind.

## References

- H. de Weerd, R. Verbrugge, and B. Verheij. How much does it help to know what she knows you know? an agent-based simulation study. *Artificial Intelligence*, 199:67–92, 2013.
- T. Hedden and J. Zhang. What do you think i think you think?: Strategic reasoning in matrix games. *Cognition*, 85(1):1–36, 2002.
- B. Meijering, H. Van Rijn, N. Taatgen, and R. Verbrugge. I do know what you think i think: Second-order theory of mind in strategic games is not that difficult. In *Proceedings of the 33rd annual conference of the cognitive science society*, pages 2486–2491, 2011.
- D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(04):515–526, 1978.
- L. Rigoux, K. E. Stephan, K. J. Friston, and J. Daunizeau. Bayesian model selection for group studies - revisited. *Neuroimage*, 84:971–985, 2014.
- W. Yoshida, R. J. Dolan, and K. J. Friston. Game theory of mind. *Computational biology*, 4(12), 2008.