

# Strategic decision making in the Stag-Hunt game

## (Bachelorproject)

Cor Steging, s2551535, s2551535@student.rug.nl,  
Harmen de Weerd\*, Trudy Buwalda\* Burcu Arslan\*

July 16, 2016

### Abstract

Research has shown that people are capable of using theory of mind, which is the ability to reason about and understand the mental state of others. 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 needs 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 across all of the participants. However, the competitive and opportunistic 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 games, as opposed to simpler strategies that only consider the current game state and not the belief state of their opponent.

## 1 Introduction

It has long been established that human beings are able to use theory of mind (Premack and Woodruff, 1978). Theory of mind is the ability to attribute mental states to others, or in other words: the ability to reason about the reasoning of other people. For example, it is possible for a chess player to reason about the reasoning of his opponent and use this to his advantage.

---

\*University of Groningen, Department of Artificial Intelligence

In strategic games, theory of mind can therefore be seen as a strategy that influences the decision making of the player. Theory of mind differs from other strategies in that the player using theory mind decides based on what he believes to be the state of mind of his opponent. In a game of chess, if a player only makes decisions based on past experiences, he is not using theory of mind. On the other hand, if the player tries to confuse his opponent with a particular move, the player is using theory of mind since he is reasoning about the state of mind of the opponent. This is referred to as first-order theory of mind. Theory of mind can be used recursively, if the player assumes that his opponent can use theory of mind as well. In second-order theory of mind, the player will try to understand what the opponent thinks that he will do. In the chess example, the player could try to hide his true intention of leading the opponent into a trap, thus using second-order theory of mind, by reasoning about the opponent's beliefs about the player's own intentions. This can be expanded out to even higher orders of theory of mind.

Earlier research has shown that using higher-order theory of mind in strategic games is difficult and cognitively demanding due to the depth of reasoning that is required (Verbrugge and Mol, 2008). Therefore one might expect people to not use theory of mind in strategic games. However, more recent studies that tested adults on strategic games have claimed that people do use theory of mind (Goodie, Doshi, and Young, 2012). Since previous research has not reached a clear consensus, this study sets out to determine whether people actually make use of theory of mind in strategic

games. The research question of this study is therefore as follows: Do people use theory of mind for decision making in games? Furthermore, if this is not the case, what strategy do they use instead? Since it has been established that it is possible for people to use theory of mind and gain an advantage from using it (Meijering, van Rijn, Taatgen, and Verbrugge, 2011), we hypothesize that people will use theory of mind for decision making in games.

In this study participants will play multiple rounds of a simple strategy game: the Stag-Hunt game. The participants will play the game with a computer agent, which will employ three different strategies; a competitive, cooperative and opportunistic strategy. The participants will need to make decisions based on what they believe to be the strategy of the computer agent. All of the decisions that the participants make in the game will be stored and, using model selection, will be compared to models of different strategies including various orders of theory of mind. This will yield the likelihood of the players using a particular strategy. If our hypothesis is correct and people do use theory of mind in the Stag-Hunt game, the decisions that are made by the participants will show high likelihoods for the theory of mind strategies.

## 2 Method

### 2.1 The Stag-Hunt model

The game that the participants will play during the experiment is a variation on the Stag-Hunt game. The Stag-Hunt game is a trust dilemma, similar to the famous Prisoner’s Dilemma, that originated in game theory. The game is played by two hunters, whose aim it is to catch either a stag or a rabbit. The smaller rabbits can be caught by either hunter individually, whereas a stag is too large for a single hunter to catch. The stag is worth more than a rabbit, therefore it is beneficial for two hunters to cooperate in order to capture it. Since the hunters do not know what choice the other hunter will make, an attempt to capture the stag without the cooperation of the other player will result in the worst possible outcome.

**Table 1: The payoff matrix of Stag-Hunt.**

	Stag	Hare
Stag	20,20	0, 10
Hare	10, 0	10, 10

The payoff matrix for the Stag-Hunt game can be seen in Table 1, where a stag is worth 20 points and a rabbit is only worth 10 points. In this table, the rows represent the choice of the first hunter and the columns represent the choice of the latter hunter. In each cell, the left number represents the points that the first hunter gained and the second number represents the points that the second hunter gained.

From the payoff matrix one can see that, unlike the Prisoner’s Dilemma, there are two pure strategy Nash equilibria; when the hunters choose to cooperate and catch a stag, and when they both choose to defect and catch a rabbit. Therefore a player of the Stag-Hunt game needs to predict the choice of his opponent such that he can mimic that choice in order to gain the maximum amount of points. This makes the Stag-Hunt an excellent game to study theory of mind with.

In theory, the Stag-Hunt game consists of a single choice; one can either choose to cooperate with their fellow hunter in order to catch a stag or choose to catch a rabbit individually. When examining the choices that people make in a game with the use of theory of mind as our hypothesis, it is essential that the players of the game have some way of reasoning about their opponent. Therefore a model was created based on the model of Yoshida, Dolan, and Friston (2008). In this model, a single player plays as one of the hunter and a computer agent takes the role of the other hunter. The most important difference between this model and the standard Stag-Hunt game is that the player in this model has more than a single moment to reason about the state of mind of his opponent in order to decide his strategy, since a game in this model consists of several turns. The model consists of a grid, much like a chess board, on which both the hunters, the stag and the rabbit are placed as seen in Figure 1. The result is a discrete, deterministic and fully

observable game. The stag and the hunters are able to traverse the grid, whereas the rabbits remain on a fixed position. Given the structure of the grid, it is impossible for a single hunter to capture the stag, but it is possible for a single hunter to capture a rabbit. When cooperating, the hunters can capture the stag by enclosing it from the adjacent sides such that the stag can no longer move. This means that is only possible to capture the stag on a spot with two adjacent sides. For example, it is therefore impossible to capture the stag in the center of the grid, since the stag would still be able to move to an adjacent side. The hunters can catch a rabbit by simply occupying the same position as the rabbit on the grid.

The game is static, sequential and adheres to a specific order: the stag always moves first, followed by the subject and then the agent. On each turn a player may move one step in any direction (as long as it is within the grid and not occupied by the stag or the other hunter) or stay on the same square. In total, every player has a maximum of fifteen turns after which the game will end. A game is therefore finished if the stag has been captured and can no longer move, if a rabbit has been captured or if the maximum amount of moves have been made. Both players gain points from capturing the stag, while only the hunter who reaches a rabbit first can catch a rabbit. The final score of a game is equal to the maximum amount of moves, which is equal to fifteen, minus the moves that were made, plus the score gained from capturing either a rabbit or a stag. It is therefore profitable to catch a prey as quickly as possible, since each move decreases your final score.

## 2.2 The computer agents

In the experiment, the computer agent that fulfilled the role of the other hunter could employ three different strategies. The strategies were used to reason about the moves that the agents should make. For every possible move that an agent could make, a utility value was calculated that depended on the type of strategy that it employed. Once these values had been calculated, the move with the highest utility value was chosen by the agent.

### Competitive strategy

The first strategy that the agent could use was a competitive strategy that ensured the pursuit of a rabbit. This was achieved by calculating a utility value based on the length of the path between a particular move and a rabbit, where a shorter path corresponds to a higher utility value.

### Cooperative strategy

Similarly, the subjects played with an agent that used a cooperative strategy that would always chase the stag, regardless of the current scenario. Its utility value was based on the path between a move and the stag, where being closer to the stag led to a higher utility.

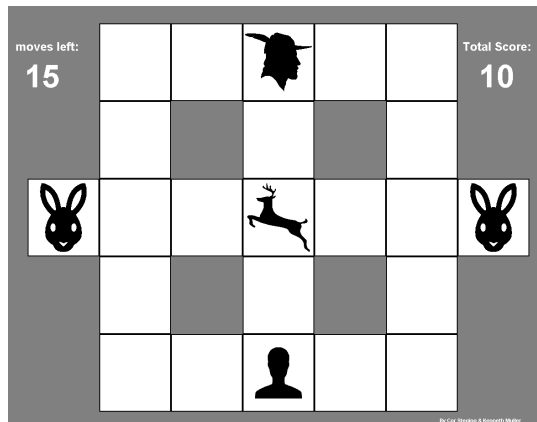
### Opportunistic strategy

Lastly, an opportunistic strategy was implemented, in which the agent chose to hunt for a stag if and only if it was close enough to it. If the stag was too far away, it would employ the competitive strategy. By using the trial and error method it was discovered that if the agent was four steps away from the the stag or closer, the cooperative strategy was profitable. Therefore the opportunistic strategy used this value to determine between defecting or cooperating.

The stag employed a single strategy that based its utility on the locations of the subject and the agent. For every possible move, the path between the stag and both of the hunters is measured. The sum of the length of these paths is used as the utility, where being further away from the hunters results in a higher utility. This ensures that the stag always moved away from the hunters.

## 2.3 The experiment

In the experiment, 15 participants each played 320 rounds of the stag hunt game on a computer. They were given the rules of the game with the objective to simply score as many points as possible. At the end of each round, the score of the player of that round was added to the total score. The total score was displayed throughout the game in order to motivate the participants. The strategy that the agent employed changed every ten rounds throughout the experiment. The participants were told that they



**Figure 1: The stag hunt game**

The subject (bottom), stag (middle) and agent (top) can move to adjacent white squares. The position of the rabbits (left and right) remain fixed throughout the game. On the left the remaining moves are presented to the user and on the right the total score is shown.

would play against agents with different strategies, however, they were not told when the strategies of the agent were changed. It was purposefully decided that the participants were not shown the scores and total scores of the agents, as this might create an adversarial urge to ‘beat’ the other hunter. During the experiment, for each round, all of the moves that were made by the participant, the agent and the stag, were recorded, as well as their initial positions.

## 2.4 Model creation

The data of the experiment can be transformed into a long list of game states and decisions. A game state consists of the positions of the stag, the agent and the participant as well as an indicator to determine whose turn it is. The decisions consists of the move made by the player in that particular state. To determine the strategies used by the participants, models of several strategies need to be created to compare the participants’ data to. These models calculate the probability of a decision for a given state, depending on the strategy of the model. The data of the participant was compared to a large number of models, however, models that showed very little compatibility with the data were excluded in order to create a clearer overview of the model likelihoods.

Most of the models that we compared the data to were created by de Weerd, Verbrugge, and Verheij (2013), unless otherwise indicated, in previous studies on theory of mind. These include three theory of mind models that mimic zero to second order theory of mind. The zero-order theory of mind model is unable to reason about the opponents state of mind and updates its own belief state only from the current game state. It uses game states from the past to predict the best possible move in the current and future states. The first-order theory of mind model can reason about the state of mind of the opponent and assumes that the opponent is a zero-order theory of mind agent. Furthermore, a first-order theory of mind agent believes that other players may change their behavior in reaction to observed behavior of others. For example, the participants might make a move in order to direct the other hunter towards the stag. Similarly, the second-order theory of mind agent reasons about the state of mind of the opponent by assuming that the opponent is a first order theory of mind agent. An example of this would be a participant making a move such that the other hunter understands that the participant wants to capture the stag.

Since it is possible for participants of our experiment to simply click random buttons in order to speed up the game, it was necessary to include a model that mimics this random behavior. This random model simply assigns the same probability to each possible move in a given game state. Two strategies based on repetition were included as well. The sticky model increases the probability of a given move if it was done previously in the same state. The win-stay, lose-shift model only increases the probability of a move if that move in the same state previously resulted in a win. If the move led to a loss, the probability of the move is decreased and the probabilities of the other possible moves in that state are increased.

Lastly, since it is possible that the participants used strategies similar to the AI opponent’s strategies, we included models of these three strategies as well. For the competitive model, the probability of each move is calculated based on the length of the path to the rabbit, where moves resulting in a shorter path are chosen with a higher probability. Similarly, the cooperative model calculates the distance to the stag for each possible move and assigns higher probabilities to the moves that lead to shorter distances. The opportunistic strategy determines the distance between the current position and the position of the stag and, if this distance is less or equal to four moves, calculates the probability for each move based on the cooperative model. If the distance is not short enough, the probabilities are calculated using the competitive model.

## 2.5 Random-effect Bayesian model selection

The method that is used to compare the participant data to the models is random-effect Bayesian model selection (Rigoux, Stephan, Friston, and Daunizeau (2014) and Daunizeau, Adam, and Rigoux (2014)). Random-effect Bayesian model selection (RFX-BMS) determines the frequency of each of the nine strategy model in the data of the participants. In other words, RFX-BMS can infer how likely it is that a given participant used each of the strategies. The models that the data is compared to are considered random effects that might be different

across participants. That is, RFX-BMS does not assume a homogeneous population, but a population in which different participants may use different strategies.

## 3 Results

Table 2 summarizes the performance of participants in the experiment. The first row represents the results across all games and the latter three rows represent the games that were played with the respective agents. We can see that across all games, the mean score per game of the participants was 21.04. With the cooperative agent the participants scored more points on average, while they scored less points with the competitive and opportunistic agents. Furthermore we can see that the stag was caught in more than half of the games with a cooperative agent and almost never with the competitive agent. In theory, the competitive agent should not be able to capture the stag, however, on rare occasions the agent can accidentally trap the stag while pursuing a rabbit.

In the games with the opportunistic agent the stag was caught 15.73% of the time. We can see that, on average, more moves were made in the games with the cooperative agent compared to the other two strategies. Since capturing a stag takes more moves and the stag was caught most of the time in the games with the cooperative agent, it makes sense that more moves were made in these games. In the next two columns of Table 2 we can see the average final score gained from capturing a rabbit and capturing a stag respectively. Overall, capturing a stag was worth almost 6 points more than capturing a rabbit, which means that even though it takes more moves to capture a stag, it is still profitable.

The agent managed to capture a rabbit in 23.46% of all games, meaning that the participants did not capture anything in these games. This did not occur in games with a cooperative agent, but it did happen in over 30% of the games with the competitive and opportunistic agents. In these games, the participants chose to cooperate in order to catch the stag, whereas the agent chose to go for a rabbit. In 5.79% of the games with

**Table 2: General results from the experiment.**

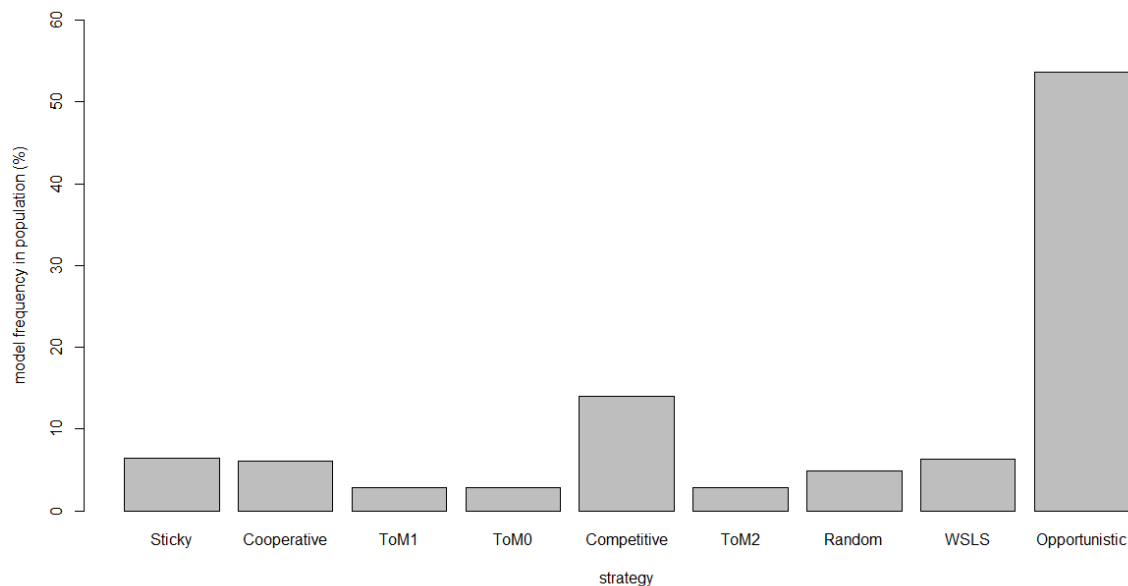
	Player score	Stags caught (%)	Moves made	Score with rabbit	Score with stag	Agent caught rabbit (%)
All	21.04	26.48	5.08	22.29	28.25	23.46
Competitive	19.7372	0.43	3.18	23.09	33	31.29
Cooperative	23.82	66.18	8.23	19.09	27.89	0
Opportunistic	19.77	15.73	4.05	22.90	29.53	37.44

the cooperative agent the participants and the agent failed to capture a prey, which meant that they were awarded zero points for that round. The data from these games are included in Table 2. In the games with the other two agents instances of players failing to capture a prey did not occur.

The result of the comparison between the decisions made by the participants and the strategy models using RFX-BMS can be seen in Figure 2. The x-axis of the bar chart represents the strategy

models and the y-axis represents the frequency with which these models prevail in the data of the participants. It should be noted that the y-axis of the chart is truncated to provide a clearer overview of the model likelihoods.

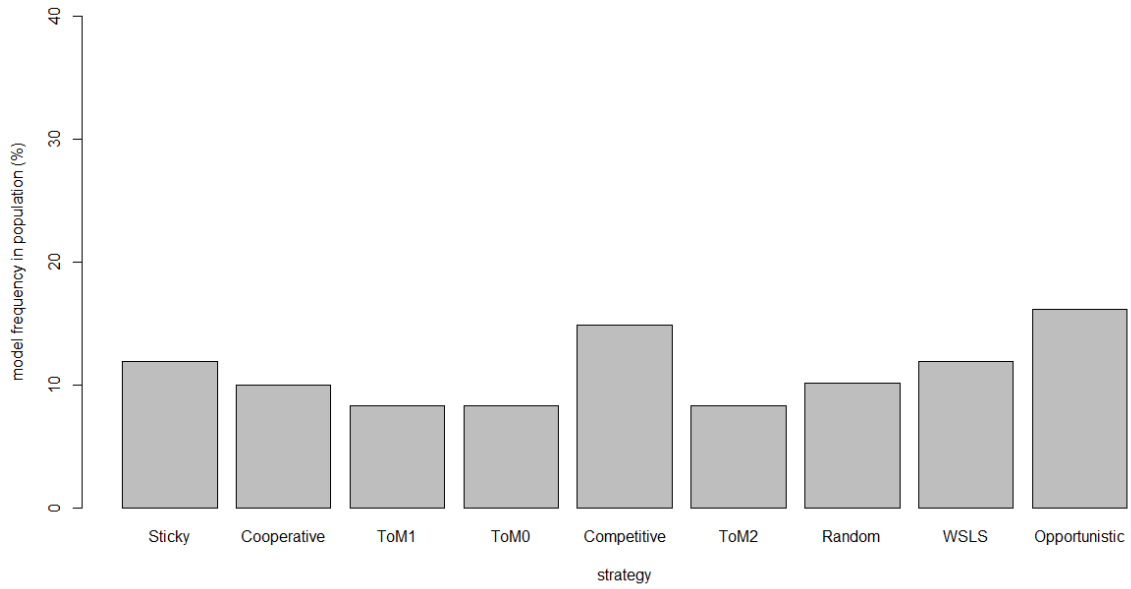
In this chart we can see that the zero, first and second-order theory of mind models (ToM0, ToM1, ToM2) have the lowest frequency across the data of the participants. This means that the theory of mind models are the least likely strategies that



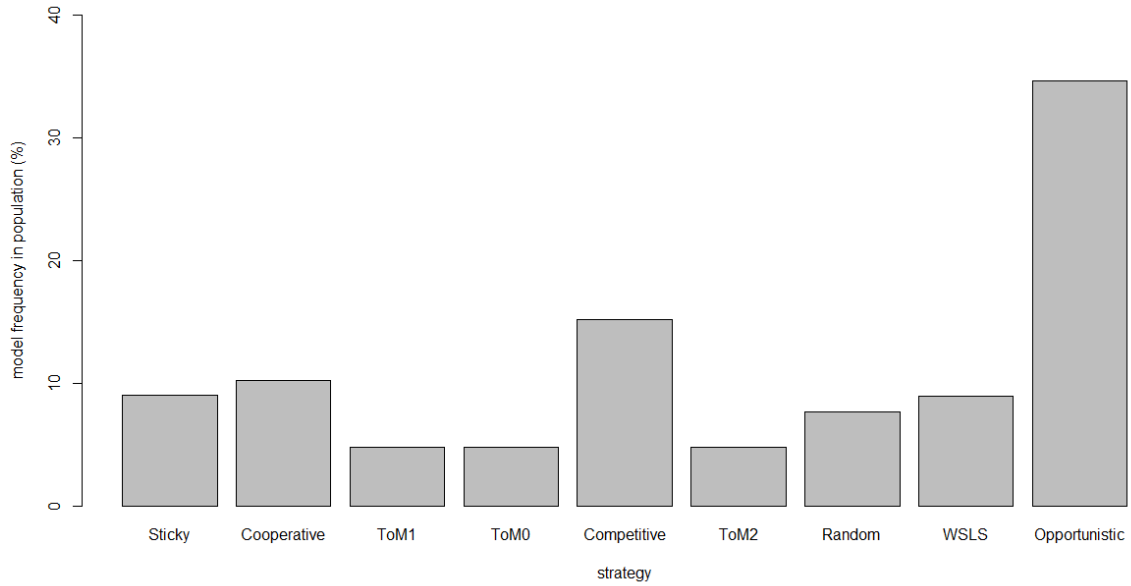
**Figure 2: The model frequency across all participants.**

**Table 3: Model frequencies of the participants per strategy of the agent in percentages.**

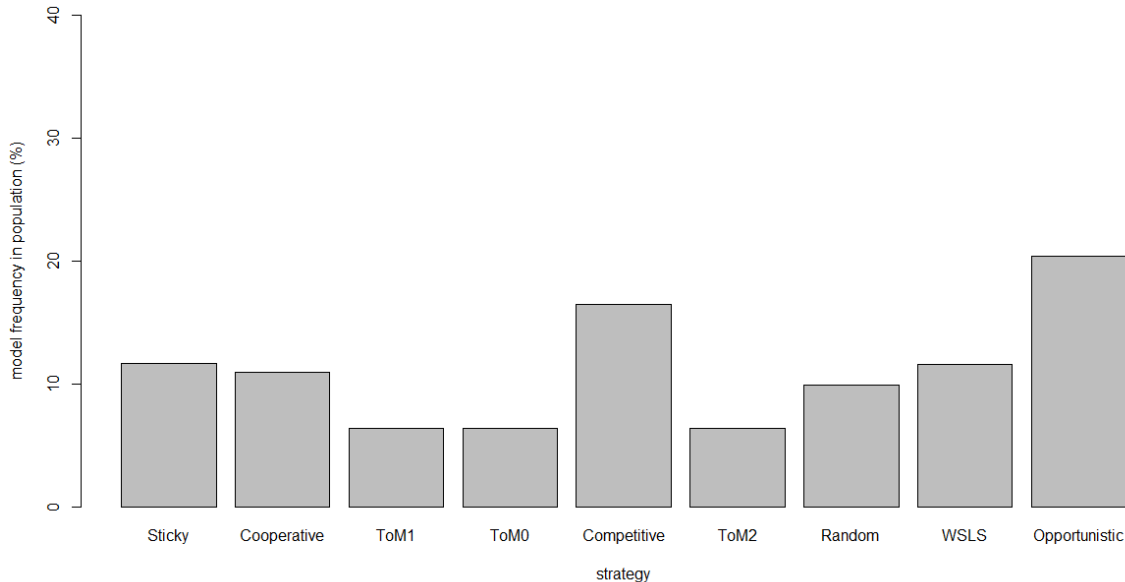
	Sticky	Cooperative	ToM0	ToM1	Competitive	ToM2	Random	WS-LS	Opportunistic
All	6.391361	6.092962	2.851491	2.851491	14.05636	2.851491	4.92365	6.335379	53.64581
Competitive	11.91498	9.967927	8.332735	8.332735	14.89828	8.332735	10.15511	11.89424	16.17126
Cooperative	9.002029	10.19757	4.776875	4.776875	15.21482	4.776875	7.631176	8.93659	34.68718
Opportunistic	11.64285	10.91049	6.361406	6.361406	16.46176	6.361406	9.906297	11.58326	20.41113



**Figure 3: Model frequencies across all participants in games with a competitive agent.**



**Figure 4: Model frequencies across all participants in games with a cooperative agent.**



**Figure 5: Model frequencies across all participants in games with an opportunistic agent.**

the participants used. The opportunistic strategy model has the highest frequency of 53.65%, followed by the competitive strategy model with a frequency of 14.06%, meaning that the participants most likely used an opportunistic strategy. The remaining strategy models, sticky, cooperative, random and Win-Stay, Lose-Shift (WSLS), all prevail across the data of the participants with a frequency of around 6%.

Table 3 contains the frequencies of each strategy model for all participants, as well as the frequencies of the strategy models split up per strategy that the agent (the other hunter) used. The rows represent the strategies of the agent and the columns the frequency of each strategy model. In this table one can see what strategies the participants employed against each strategy of the agent. In Figure 3, 4 and 5, the model frequencies of the participants against each strategy of the agent are plotted in three different bar charts.

As was discussed in Section 2.1, the optimal strategy of a player is to mimic the strategy of his opponent. Therefore, when playing with a competi-

tive agent for example, one might expect to see the competitive strategy model prevail. In Figure 3, 4 and 5 and Table 3, however, we can see that this is not the case. With the exception of the Opportunistic strategy, which prevailed across all games, there is no reason to suggest that the participants used the same strategy as the agent. Instead, we can see that the participants are most likely to use the opportunistic strategy against all of the strategies of the agent. Furthermore, we can see that the differences in frequency between each strategy is larger when playing with a cooperative or opportunistic agent as opposed to playing with a competitive agent.

## 4 Discussion

The decisions that the participants made in the Stag-Hunt game were gathered and compared to strategy models in order to understand what strategies participants used. We hypothesized that people use theory of mind in decision making, however, the results of this study do not support that hypothesis. In Figure 2 and Table 3 we can clearly see that all of theory of mind models prevail



with the lowest frequency across all participants. The opportunistic model, however, prevails with an exceptionally high frequency, which suggests that the participants used the opportunistic strategy. In this strategy, participants would only hunt for the stag if the stag was within four moves away from the participant, otherwise the participants would hunt for a rabbit.

As we can see in Figure 3, 4 and 5, the participants did not use the same strategy as their fellow hunter, which would have been the optimal strategy. This could imply that the participants did not reason about the strategy or state of mind of the opponent, providing more evidence to suggest that the participants did not use theory of mind in this game. Interestingly, the RFX-BMS results with the cooperative agent are much clearer than the results with the competitive agent. This could be due to the fact that it is almost impossible to catch a stag with a competitive hunter. If a participant set out to catch a stag, but the competitive agent did not cooperate, the participant would have to change his strategy half-way into a game, which would have led to a mixed strategy and therefore unclear results. The same phenomenon could occur with an opportunistic agent, since the opportunistic strategy essentially becomes a competitive strategy if the stag is too far away from the agent. Therefore if the participant initially wants to catch the stag, but the stag is too far away from the agent, the participant might change his strategy if he notices that the agent is not cooperating. This could explain why the results with the opportunistic strategy are less clear than those with the cooperative agent and clearer than those with the competitive agent.

Some factors could have influenced the results gained from the experiment. First of all, the participants may have shown a learning effect. After the experiment ended, some participants informally mentioned that they based their strategy solely on the first move of the agent. The result of this is that the participants stick with a single profitable strategy instead of adapting their strategy based on the perceived state of mind of the opponent. Fatigue or a lack of concentration might have played a role in the strategies of the participants as well. The participants had to play 320 trials of

the same game, which in total took approximately thirty minutes. Since the participants were told how long the experiment would take ahead of time, it is possible that they were not as concentrated on the game. This could have led to the participants making simple decisions that do not use theory of mind in order to speed up the game, which could explain the results that were found.

The Stag-Hunt game exposes the participants to extra variables as well, which could have influenced the results. In the Stag-Hunt game, the participant does not only need to concentrate on the moves that the opponent makes, but also on the decisions and moves that the stag makes. This extra variable could influence the decision making of the participants. Lastly, the agent that fulfilled the role of the other hunter employed three relatively simple strategies. Perhaps if the agent would have used some order of theory of mind, the participants might have identified more with the agent's decisions and have used theory of mind as well.

#### 4.1 Future research

In future research one could expand on this study by altering the settings of the experiment. A different game, preferably one with less variables than the Stag-Hunt game, could be chosen to perform the research with, to determine whether the game itself influences the results. Furthermore the strategies of the agents could be expanded upon such that the agents themselves use theory of mind. Another variable in this study that could be changed in further research is the frequency with which the agent changes strategies. Perhaps participants are more likely to use theory of mind if the agent maintains a strategy for a longer period of time.

## 5 Conclusion

To conclude, this study discovered that people are not likely to use theory of mind in the Stag-Hunt game. Instead people are more likely to use an opportunistic strategy that does not make use of the state of mind of the opponent.

## References

- J. Daunizeau, V. Adam, and L. Rigoux. Vba: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Computational Biology*, 10, 2014.
- 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*, volume 199-200:pages 67–92, 2013.
- A.S. Goodie, P. Doshi, and D.L. Young. Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*, volume 25(1):pages 95–108, 2012.
- B. Meijering, H. van Rijn, N.A. 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*, volume 85(1):pages 2486–2491, 2011.
- D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, volume 1(4):pages 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.
- R. Verbrugge and L. Mol. Learning to apply theory of mind. *Journal of Logic, Language and Information*, volume 17(4):pages 489–511, 2008.
- Wako Yoshida, Ray J. Dolan, and Karl J. Friston. Game theory of mind. *PLoS Computational Biology*, 4:1–14, 2008.