



# ARE ONLINE GAMERS BETTER AT THEORY OF MIND REASONING THAN THE GENERAL POPULATION?

Bachelor's Project Thesis

Jouke van der Weij, s2003325, j.v.van.der.weij@student.rug.nl,

Supervisor: Prof. dr. R.L.C. Verbrugge

**Abstract:** Theory of Mind reasoning is reasoning about the mental states of other individuals. There are several orders in theory of mind reasoning. In this paper we investigate whether online gamers use theory of mind reasoning with greater success than non-gamers. Participants had to play two different games, Marble Drop and Colored Trails, where theory of mind reasoning is necessary to maximize results. This study shows that online-gamers do use theory of mind reasoning more efficiently than non-gamers and have a higher accuracy in the first game and a higher score in the second.

## 1 Introduction

In psychology theory of mind reasoning is known as reasoning about the mental states of other individuals. The ability to use theory of mind reasoning helps us both in social aspects and in decision making in critical situations (Premack, 2007). This kind of reasoning is known to develop in the early stages of childhood and it is known that children at the age of 4 can use some form of theory of mind reasoning. Astington and Edward (2010) showed that at the age of 2 children already understand that there is a difference between what they think and what is happening in the world. At the age of 4 children understand that some thoughts might not be true and what they believe to be true can be different from what others think to be true.

### 1.1 The different forms of theory of mind

One of the first papers where higher-order theory of mind reasoning was described was (Verbrugge, 2009). The basic form of theory of mind reasoning is called the zero-order theory of mind reasoning. Technically zero-order theory of mind is not a form of theory of mind reasoning but the ability to reason about world facts. It is called zero-order, because it forms the basis of higher order theory of mind reasoning. First-order theory of mind reasoning, a form used by four year old in the paper of Astington and Edward (2010), is the ability to

not only assign mental states to oneself but also to other individuals. First-order theory of mind reasoning is a form used on a daily basis by all humans over the age of 4. The ability to use first-order theory of mind makes humans able to reason about the effects of their actions to the mental states, beliefs, actions and emotional status of other individuals. Using first-order theory of mind is an important aspect in the social behaviour of humans. Reasoning about what might cheer up a friend, because he looks sad is an example of first-order theory of mind. But human cognition does not end here. Humans also can successfully use second-order theory of mind reasoning. This form of theory of mind is the ability to not only form beliefs about the mental states of other individuals, but also form beliefs about the beliefs of the other individual about your own mental states and those of others. Sarcasm for example only works when the other knows you are not serious and the user knows the other knows you are not serious and knows what you really mean. Using sarcasm requires the user to use second-order theory of mind. Meijering (2014) showed that people find it harder to make a prediction about a decision than making the exact same decision themselves. In their experiment, participants were forced to make a decision or to predict the decision of others in the exact same situation. Making a decision stood equal to "What does the opponent intend to do" and thus required first-order theory of mind, while predicting, "What does the opponent think that I intend to do?" and thus

required second-order theory of mind and an extra step. Participants took longer and made more mistakes while predicting than when deciding themselves. This suggests that performance decreases when the thought process requires an extra step. In other words, humans perform less accurately when using second-order theory of mind than first-order theory of mind.

## 1.2 Gaming and reasoning

As of today a lot of different games can be played online and involve some form of social interaction. Earlier studies by Graham (2014) showed that repeated online gaming can both have positive and negative effects on social behaviour of individuals. Graham made a distinction between prosocial and antisocial behavior. Antisocial behavior included trait hostility and the enjoyment of conflict. Positive effects included an increase of prosocial behavior in light gamers. Graham called light gamers, players who played video games less than average, but more than none at all. Prosocial behavior was measured by two variables; Trait altruism, and cooperation. Negative effects included an increase in anti-social behavior in heavy gamers, gamers who played games more than average.

Kowert and Oldmeadow (2013) showed a relationship between video game involvement and emotional sensitivity. Players who are more involved in the video game show an increased ability to identify non-verbal communication of others. In other words, they could use first-order theory of mind more successfully to determine the mental state of other individuals, than the less involved players. In most online games the player requires knowledge about the possible actions and mental states of their opponent to successfully reach their goal.

The skill of reasoning about the mental states of the opponent might seem useful in these particular games, but does not necessarily mean that these skills can be used outside the closed environment of that game. Putting the online-gamers in a different area with a different goal might show what they exactly learned. Studying the decisions of online-gamers in games other than those they are playing online, might find a difference to those people

who do not play games where human-to-human interaction is necessary to be successful in the game. By letting a group of online-gamers and a group of non-gamers play two different games, one competitive and one mixed-motive, and measuring their performance, I hope to find a difference between online-gamers and non-gamers and an answer to the question: “Does online gaming have an effect on the use of theory of mind reasoning of individuals?”

## 1.3 Hypothesis

Arslan, Verbrugge, Taatgen, and Hollebrandse (2015) showed that children at the age of 5-6 years have an increased performance in second-order theory of mind reasoning when they get feedback on their decision. Even a simple correct/wrong feedback response has a positive effect on their performance. This would suggest that feedback in the form of kill or be killed -in many games this is the ultimate goal- should stimulate the use of higher orders theory of mind. Further research in Youngblade and Dunn (1995) showed that pretend play has a strong link to the development of understanding other people’s feeling and beliefs in 40 months old siblings. As online gaming can be seen as a form of pretend play, this suggests that online gaming can have an effect on the development of similar skills.

# 2 Method

## 2.1 The participants

For this experiment the participants were divided into two groups: online-gamers and non-gamers. The online-gamers consisted of participants who have been playing online games regularly for at least one hour per week. The games played by the online-gamers were restricted to games where human to human interaction is necessary while playing the game. This includes games such as shooters (FPS), multi-massive online playing games (MMORPG) and others. The non-gamers, or the control group, are participants who have not played any online games recently. Both groups consisted of students age between 18 and 30 years old. All participants were not colour blind and spoke Dutch.

Each group consists of 15 individuals.

## 2.2 Materials

For this experiment I used a laptop with mouse and a program provided by Eveline Broers (Broers, 2014). The program consisted of a questionnaire to determine whether the participant is colourblind or not, a game called Marble Drop and a game called Colored Trails. Both games will be explained later on.

### 2.2.1 Questionnaire

The questionnaire consisted of questions as can be seen in figure A.1 and A.2 in the appendix. These questions determine whether the participant is colourblind or not and is capable of doing the experiment. After the questionnaire the experiment starts.

### 2.2.2 Marble Drop

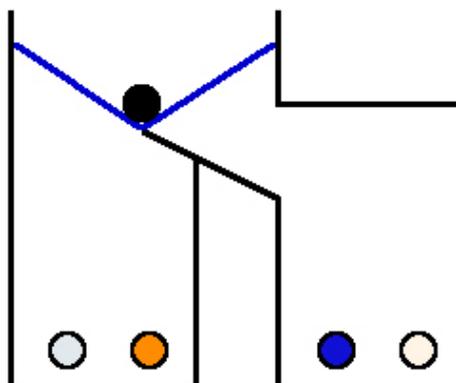
Marble Drop is a game first introduced by Meijering, van Maanen, van Rijn, and Verbrugge (2010) in their PHD thesis “Reasoning about self and others”. In the game Marble Drop, the participant is presented with a marble which he or she has to drop into the compartment which contains the marble in the darkest possible shade of his or her colour. The participant can reach the right compartment by opening a gate which the marble will fall through. The decision of the participant is correct when he or she reaches the darkest possible shade and their decision is wrong when they do not. There are three different levels of this game. In two of the three levels an opponent is presented who also controls one set of gates. This opponent, in this paper referred to as “AI\*” is a computer programmed to always make the correct decision and it will always assume the participant will make the correct decision and act upon that assumption.

In the first level the participant controls one set of gates and the AI controls no gates (Figure 2.1). In this level the participant does not need to use theory of mind reasoning to get to the correct compartment.

---

\*It’s not an actual AI, but for the convenience it will be referred to as AI.

### Example of zero order ToM Marble Drop



**Figure 2.1:** In this example the participant has the colour blue and thus should try to reach the darkest blue marble possible. The participant controls the blue gates and can open them by either clicking the left or right gate. In this case, the participant should open the right gate to reach the darkest blue marble.

The second level has an extra compartment and now the participant controls one set of gates and the AI also controls one set of gates (Figure 2.2). In these trials the participant is required to use first-order theory of mind reasoning in order to reach their goal. The last level contains another extra compartment and another set of gates is added which are controlled by the participant. Now the participant controls two set of gates and the AI still controls one, see Figure 2.3. In these trials the participant is required to use second-order theory of mind reasoning in order to reach their goal.

The goal for the opponent is the same as the participant for both levels where the opponent also controls a set of gates. The opponent also tries to get the marble in the compartment containing his darkest possible marble. Both the participant and the AI are aware of the goal of their opponent.

### 2.2.3 Colored Trails

In the second part of the experiment based on the game ‘Colored Trails’, the participant needs to negotiate with an agent about tiles with a pattern on

### Example of first order ToM Marble Drop

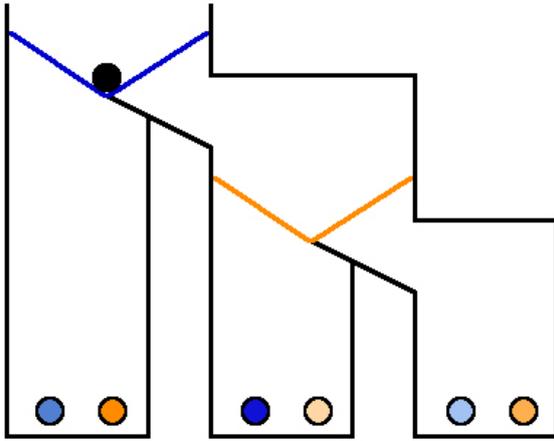


Figure 2.2: In this example the participant has the colour blue and thus should try to reach the darkest blue marble possible. The participant controls the blue gates and can open them by either clicking the left or right gate. The AI controls one set of orange gates. In this case, the participant should open the right gate to reach the darkest blue marble.

### Example of second order ToM Marble Drop

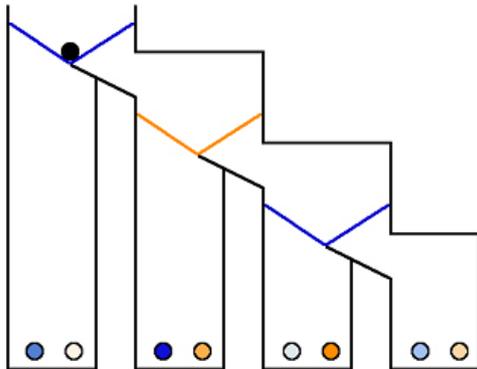


Figure 2.3: In this example the participant has the colour blue and thus should try to reach the darkest blue marble possible. The participant controls the blue gates and can open them by either clicking the left or right gate. The AI controls one set of orange gates. In this case the participant should open the right gate as his opponent will open his left gate, because the opponent thinks the participant will open his right gate after that.

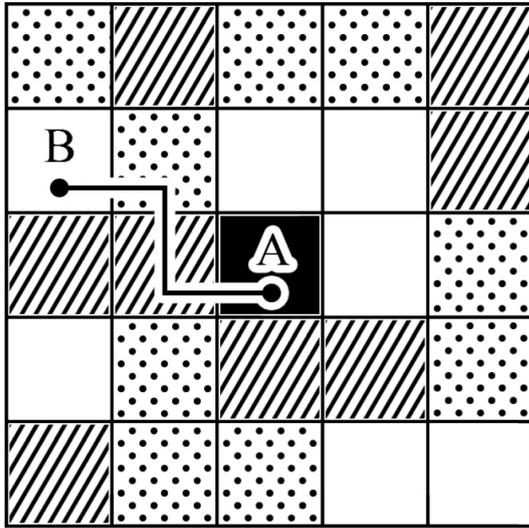
them which are required to reach their goal. The participant can reach their goal by matching the patterns on the tiles in their hand with the patterns on tiles between their start position and their goal. Appendix A.3 shows the screen in which Colored Trails takes place. The playing field consists of a board made out of 5 by 5 squares. The start position of the participant is always at the center of the field. The goal, indicated by a tile with a blue border, is always 3 or more tiles away from the start position. Steps are made in either horizontal or vertical direction, never diagonal. The goal can only be at the darkened positions in Figure 2.4.

Furthermore, the participant does not know the goal of his opponent and the opponent does not know the participant's goal. Each trial lasts at most six rounds, alternating turns between the participant and the agent. In each round the players have at most one minute to make a decision. On the participant's turn, he or she can either choose to accept the offer he or she received, make an offer, or stop the negotiations. When the participant stops the negotiations, the starting hands of both players is used to calculate the score. At the turns of the agent, the agent can choose between the same actions as the participant.

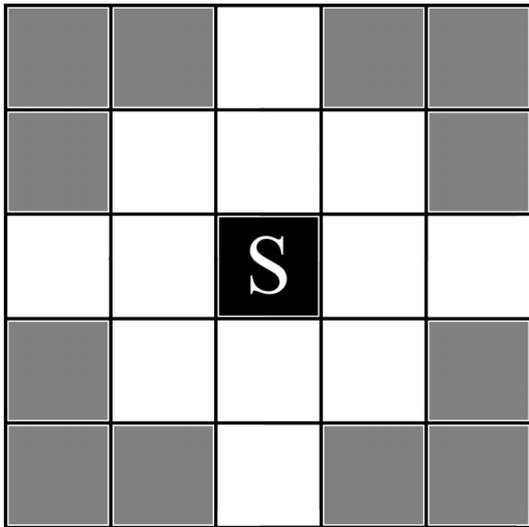
The goal of Colored Trails for a player was to get as many points as possible. Each trial the player started with 50 points. When the player reached the goal it got an extra 50 points. Each tile left over after reaching the goal was worth another 5 extra points. When the player did not reach the goal, each step still needed to reach to goal costed the player 10 points. After each trial the end score was added to the total score of the previous trials.

### Agents

There are three agents in Colored Trails each using a different order of theory of mind reasoning. The so called ToM<sub>0</sub>, ToM<sub>1</sub> and ToM<sub>2</sub> (De Weerd, Verbrugge, and Verheij, 2013) each use theory of mind reasoning an order higher than their previous agent, starting at zero-order theory of mind reasoning and ending at second-order theory of mind. In the PHD thesis of de Weerd (2015), ToM<sub>0</sub> is described as "A ToM<sub>0</sub> agent is essentially fixated on his own piece of pie, and does not consider the piece of pie of other agents at all. Instead, the



(a)



(b)

Figure 2.4: Board (a) shows a possible route from the start position, A, to the position B. Here the player needs one diagonal striped tile, one dotted tile and a blank tile to reach Position B. Board (b) shows all possible locations for the goal position presented by the gray tiles and the start position presented by the letter ‘S’. These pictures are taken from de Weerd (2015) with the permission of H. de Weerd.

ToM<sub>0</sub> agent constructs zero-order beliefs about the likelihood that his trading partner will accept a certain offer” (de Weerd, 2015). This essentially means that the ToM<sub>0</sub> agent uses no theory of mind reasoning and only reasons about the behavior part of the other agent to decide what action to take, relying on the observations the agent made.

The ToM<sub>1</sub> agent is described in the same PHD thesis as “The use of first-order theory of mind allows a ToM<sub>1</sub> agent to put himself in the position of his trading partner to consider an offer from the perspective of his trading partner”. This means that the agent can make beliefs about the goal of the participant and calculate a probability based upon those beliefs that a participant accept certain offers, thus essentially using first-order theory of mind reasoning.

The ToM<sub>2</sub> agent uses second-order theory of mind reasoning and besides from reasoning about the possible goals of the participant, also considers the beliefs of the participant about the agent’s goal and beliefs (de Weerd, 2015).

#### 2.2.4 Statistical tests

During Marble Drop, the answer of the participant was recorded for each trial. For each trial it was recorded whether the participant was correct or not. By dividing the correct answers per level by the total trials per level, an accuracy was calculated for each level for each participant. During Colored Trails, for each participant the points gained per trial and the total points gained per agent was recorded. Both the data gained from Marble Drop and Colored trails can then be compared between the two groups using a t-test.

### 2.3 The procedure

The experiment starts with the questionnaire containing 10 questions in the form of the questions in Appendix A.1. If all the questions are answered correctly the experiment proceeds with 10 questions in the form of the questions in Appendix A.2. If all of these questions are answered correctly the actual experiment starts.

### 2.3.1 Marble Drop

The first game of the experiment was Marble Drop. At first the participant got to read a detailed description of the game, explaining the rules and goal of the game. Marble Drop consisted of a total of 20 trials. the first four trials were zero-order theory of mind trials. After these the experiment continued with 8 trials of first-order theory of mind Marble Drop. The experiment ended with 8 trials of second-order theory of mind Marble Drop. After each trial in each level the participant got feedback on their answer. If the answer was correct the participant got the message “Correct”. If the participant was wrong, the participant saw a message “Wrong. You could have ended here” with an arrow pointing to the correct compartment. After each level the participant got feedback on how many trials were correct out of the total number of trials. After all 20 trials of Marble Drop were played the second game, Colored Trails, started.

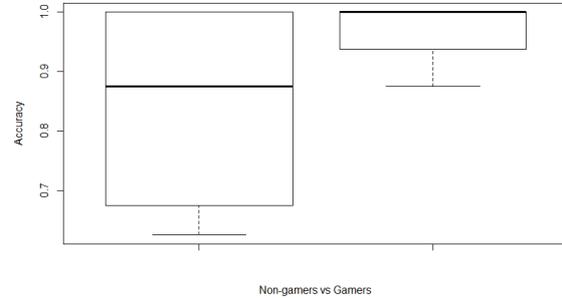
### 2.3.2 Colored Trails

The second game of the experiment was Colored trails. This session starts with a very detailed description of the rules and goals of the game. After the description a short questionnaire had to be answered to make sure the participant understood the game. The game itself consisted of 24 trials, each lasting at most 6 rounds. Each participant played 3 sessions of 8 trials each. Each session was against a different agent mentioned in the materials section. The order in which the participants played against each agent was random for each participant. After each trial, the participant saw how many points he or she gained during the trial. After the third session the participant saw his or her total score and got the message that the experiment was over.

## 3 Results

All participants had a perfect score in the first level of Marble Drop, meaning a ceiling effect is reached in zero-order theory of mind in Marble Drop. Thus these results will not be further discussed in this paper.

### The difference in accuracy between the two groups in first-order Marble Drop



**Figure 3.1:** The results in accuracy for first order marble drop in non-gamers (left) and online-gamers (right).

### 3.1 Marble Drop

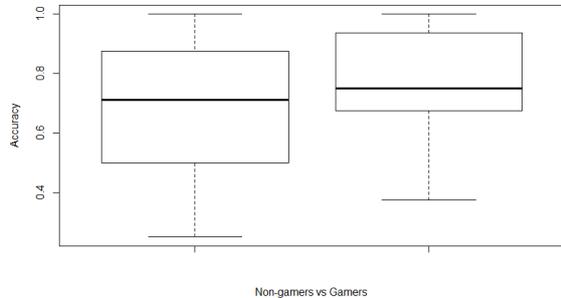
Figure 3.1 shows the results of the two groups in the second version of Marble Drop, where first-order theory of mind was needed. Both groups score relatively high with a mean of 0.87 for non-gamers and 0.97 for online gamers. For online gamers, this possibly means another ceiling effect is effective when first-order theory of mind is required. This is further suggested by the median of 1.0 for online gamers. For non-gamers, the median is 0.875. This is related to 8 out of 8 decisions and 7 out of 8 decisions correct, respectively.

Figure 3.2 shows the results of the two groups in the last version of Marble Drop, where second-order theory of mind was needed. This also shows a difference in accuracy between non-gamers and online-gamers, although the difference is slightly less than the difference between them in first-order theory of mind. Both groups score a bit lower than in the first order Marble Drop with a mean of 0.72 for non-gamers and 0.78 for online-gamers. The median for the two groups is 0.71 and 0.75, respectively. Overall, the online-gamers thus have a slightly higher accuracy than the non-gamers.

### 3.2 Colored Trails

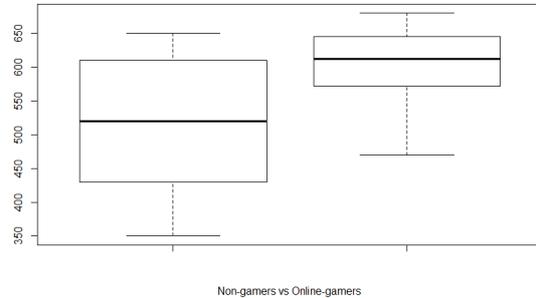
First, looking at the total scores shows a difference between the two groups. Figure 3.3 shows the to-

**The difference in accuracy between the two groups in second-order Marble Drop**



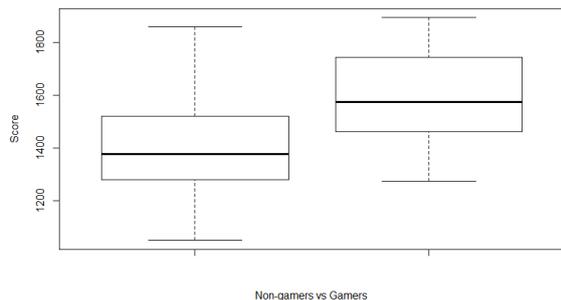
**Figure 3.2:** The results in accuracy for second order marble drop in non-gamers (left) and online-gamers (right).

**The difference in total scores against a Tom<sub>0</sub>**



**Figure 3.4:** The total scores of the non-gamers (left) and online-gamers(right) against a zero order theory of mind agent

**The difference in score between the two groups in Colored Trails**



**Figure 3.3:** The results in score for second order marble drop in non-gamers (left) and online-gamers (right).

tal scores of all the participants. It shows a difference between the scores of the two groups, with the online-gamers having a higher total score than the non-gamers. The means of the two groups are 1590 and 1401, respectively. The median of the two groups are 1575 for online-gamers and 1378 for non-gamers.

Looking at the results per trial, we see a similar trend. The histogram in Appendix A.4 shows the frequency of each score for each trial. Appendix A.4 shows that there is a difference in frequency of the 100-score. The online-gamers group

has significantly more trials where they scored 100 points than the non-gamers group, indicating that online-gamers reach their goal more often than non-gamers. As a result the frequency of the lower scores is higher in the non-gamers group. The frequencies of the scores higher than 100 are about equal in both group but such high scores do not occur very frequently.

**3.2.1 Results per agent**

As of now, the results do not show the difference per type of opponent, but it might be interesting to look at the difference between the two groups per zero-, first- and second-order agent. Figure 3.4 shows the difference between the two groups in the total scores against a zero-order theory of mind agent.

Figure 3.4 shows that there is a difference between the two groups of participants, with the major scores of the non-gamers ranging between 425 and 610, and the scores of the online-gamers ranging between 575 and 650. The means of the two groups are 518 and 604, respectively. Appendix A.5 shows an histogram of the scores per trial against a ToM<sub>0</sub> agent. This histogram shows a similar trend as the total scores overall. The frequency of reaching 100 points is higher in the online-gamers, indicating that the online-gamers reach their goal more often. The same data is plotted in Figure 3.5 and Appendix A.6 for the ToM<sub>1</sub> trials and in Figure 3.6 and Appendix A.7 for ToM<sub>2</sub> trials.

The difference in total scores against a  $Tom_1$

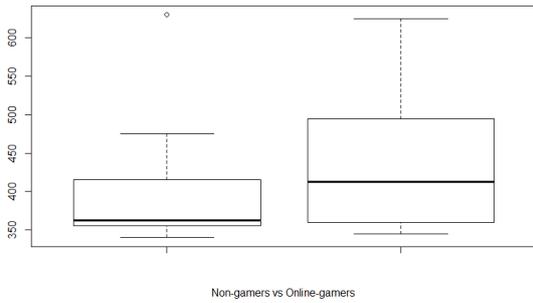


Figure 3.5: The total scores of the non-gamers (left) and online-gamers(right) against a zero order theory of mind agent

The difference in total scores against a  $Tom_2$

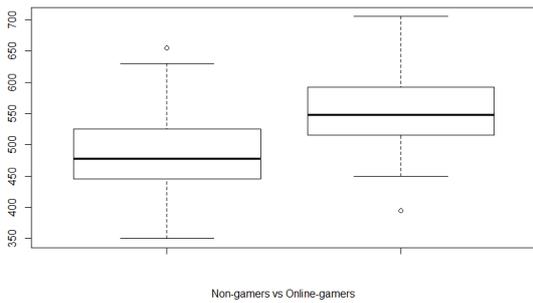


Figure 3.6: The total scores of the non-gamers (left) and online-gamers(right) against a zero order theory of mind agent

For both  $Tom_1$  and  $Tom_2$  trials, a similar trend can be seen as before. The online gamers have a higher frequency of reaching 100 points, indicating that they reach their goal more often.

### 3.3 Statistical Tests

To check whether there is a difference in the performance in the two games, multiple t-test are performed. A t-test can test whether the mean of a value differs between two groups. For Marble Drop a t-test is performed for each session. This t-test is used to compare the mean accuracy of the non-gamers and online gamers for the zero-order theory of mind trials. This is repeated for the first-order theory of mind trials and the second-order theory of mind trials. Another couple of t-tests can show the difference in scores for non-gamers and online-gamers against all three different types of agents.

#### 3.3.1 Marble Drop

In the zero-order theory of mind trials, both groups had a perfect performance, thus no significant difference is found. As a ceiling effect is present here, no t-test is needed. In the first-order theory of mind trials, a t-test gives a p-value of 0.02, so a significant difference is found between the two groups in first-order theory of mind Marble Drop. This means that online-gamers have a significantly higher accuracy than the non-gamers in this session. In second-order theory of mind, a t-test gives a p-value of 0.047, thus a significant difference is found between the two groups in second-order theory of mind Marble Drop. This means that online-gamers have a significantly higher accuracy than the non-gamers in this session.

#### Colored Trails

A t-test performed for the scores against a zero-order theory of mind agent gives a p-value of 0.01, thus a significant difference is found between the two groups against a zero-order theory of mind agent. This means that online-gamers scored significantly higher points against a zero-order theory of mind agent than non-gamers. A t-test performed for the scores against a first-order theory of mind, gives a p-value of 0.18, thus no significant difference is found between the two groups against a

first-order theory of mind agent. This means that there is no significant difference in score between the online-gamers and the non-gamers against a first-order theory of mind agent. A t-test performed for the scores against a second order theory of mind, gives a p-value of 0.058, thus no significant difference is found between the two groups against a second order theory of mind agent. This means that there is no significant difference in score between the online-gamers and the non-gamers against a first-order theory of mind agent

## 4 Conclusion and Discussion

Looking at the results, we can make two different conclusions out of the statistical tests. There is a significant difference between the two groups in the ToM<sub>1</sub> and ToM<sub>2</sub> versions of Marble Drop. Gamers overall have a higher accuracy in both versions, indicating that they do have an increased performance when using theory of mind reasoning than non-gamers. Note that in Marble Drop the decision of the AI is always clear when using theory of mind reasoning. This is because the goal of the AI is always visible, and there is only one way to reach that. This makes theory of mind reasoning more certain to be correct.

When looking at the results of Colored trails, we do see a higher score of the gamers than the non-gamers, but this difference shows to be not significant. This would suggest that there is no increased performance in theory of mind reasoning of the online-gamers. This at first hand looks contradictory to what was found in the results of Marble Drop. This can be due to the fact that while using higher-order theory of mind reasoning, a gamer comes to the wrong decision of what might be the goal of the agent. but this is contradictory to the results found in the paper of de Weerd et al. They found that using higher order theory of mind against a higher-order theory of mind agent results in a higher score de Weerd (2015). The difference in conclusion found for both games is up for discussion which might require more research. An other reason could be that the two groups were too small to prove a significant difference. With 15 members per group the power of the statistical tests are relatively low.

Earlier research found that individuals have a hard time reaching pareto optimal in Colored Trails (Broers, 2014). Eventually for each situation there is an optimal decision which maximizes the score obtained. More blocks against each agent might result in online-gamers eventually reaching pareto optimal more often, or sooner than the non-gamers. Another explanation could be that in most games the gamers are used to, the situations wherein theory of mind reasoning can be used are competitive situations, they are not cooperative or do not require negotiation, thus the theory of mind learned during those games can only be transferred to competitive situations such as Marble Drop.

### 4.1 Future Work

For future work, the data can be used to make prediction about what theory of mind the participants used in Colored Trails. By using a ToM<sub>3</sub> agent, described in (Weerd, Verbrugge, and Verheij, 2015). This agent is a ‘spectator’ and reasons based on the offers of the participants what theory of mind the participant used. This agent produces a confidence of how similar the offers of the participant were to those of the ToM<sub>0</sub>, ToM<sub>1</sub> or ToM<sub>2</sub> agent. A higher confidence in one order theory of mind than the confidence of an other theory of mind, means it was more likely that the participant used that order theory of mind.

An other method to determine the strategy used by the participants is by comparing the likelihood of different strategies using Bayesian model comparison described by (Stephan, Penny, Daunizeau, Moran, and Friston, 2009). This method uses the data to compile a list of probabilities for each strategy used in the comparison. By using the obtained data in this experiment and the Bayesian model comparison, we can conclude what strategy each participant used. Comparing the strategies used by the participants in two groups, might further explain if theory of mind was a more dominant strategy in online-gamers than non-gamers.

## 5 Acknowledge

I would like to thank E. Broers for providing the code needed to run the experiments. I would like to thank my supervisor dr. L.C. Verbrugge for the guidance she offered during the process. Furthermore, I would like to thank H. de Weerd for the availability to ask questions and answering them.

## References

- Burcu Arslan, Rineke Verbrugge, Niels Taatgen, and Bartjan Hollebrandse. Teaching children to attribute second-order false beliefs: A training study with feedback. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, pages 108–113, 2015.
- Janet Wilde Astington and Margaret J Edward. The development of theory of mind in early childhood. *Social Cognition in Infancy\** 5, page 16, 2010.
- E. Broers. Negotiating with incomplete information: The influence of theory of mind. Master’s thesis, University of Groningen, The Netherlands, 2014.
- Harmen De Weerd, Rineke Verbrugge, and Bart Verheij. How much does it help to know what she knows you know? an agent-based simulation study. *Artificial Intelligence*, 199:67–92, 2013.
- Hermanes Albertus de Weerd. *If you know what I mean: Agent-based models for understanding the function of higher-order theory of mind*. PhD thesis, 2015.
- Jeffrey Graham. Video games: Parents’ perceptions, role of social media and effects on behavior. In *Media and Communications– Technologies, Policies and Challenges*. Nova Science Publishers, Inc, 2014. ISBN 9781633210158.
- Rachel Kowert and Julian A. Oldmeadow. (a)social reputation: Exploring the relationship between online video game involvement and social competence. *Computers in Human Behavior*, 29(4): 1872 – 1878, 2013. ISSN 0747-5632.
- B. Meijering, L. van Maanen, H. van Rijn, and R. Verbrugge. The facilitative effect of context on second-order social reasoning. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pages 1423–1428, 2010.
- Ben Meijering. *Reasoning about self and others*. PhD thesis, 2014.
- David Premack. Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences*, 104(35):13861–13867, 2007.
- Klaas Enno Stephan, Will D Penny, Jean Daunizeau, Rosalyn J Moran, and Karl J Friston. Bayesian model selection for group studies. *Neuroimage*, 46(4):1004–1017, 2009.
- Rineke Verbrugge. Logic and social cognition. *Journal of Philosophical Logic*, 38(6):649–680, 2009.
- Harmen Weerd, Rineke Verbrugge, and Bart Verheij. Negotiating with other minds: the role of recursive theory of mind in negotiation with incomplete information. *Autonomous Agents and Multi-Agent Systems*, pages 1–38, 2015.
- Lise M Youngblade and Judy Dunn. Individual differences in young children’s pretend play with mother and sibling: Links to relationships and understanding of other people’s feelings and beliefs. *Child Development*, 66(5):1472–1492, 1995.

## A Appendix

**Vraag 1. Welke van deze twee vlakken is blauw?**

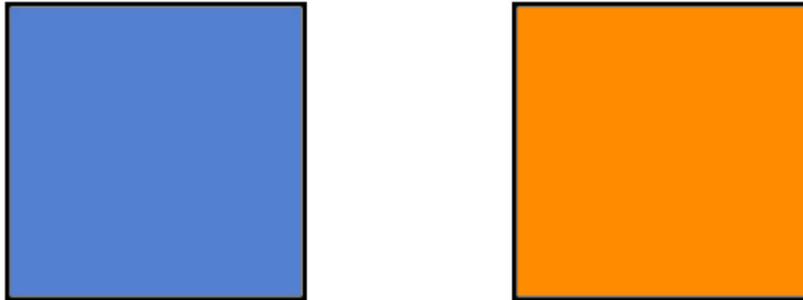


Figure A.1: Example question: The participants needs to press the square in the colour asked. The question above reads: “Which of the two squares is blue?”

**Vraag 1. Welke van deze twee vlakken is lichter?**

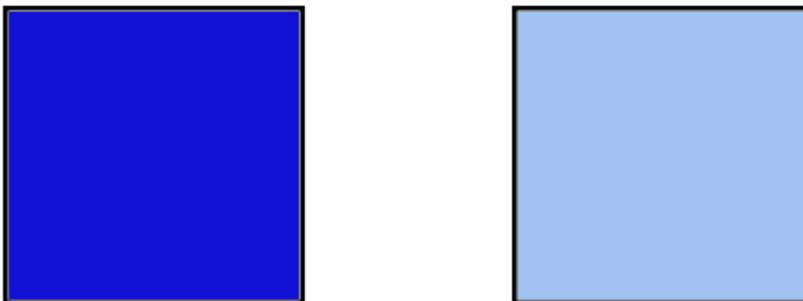


Figure A.2: Example question: The participants needs to press the square in the colour asked. The question above reads: “Which of the two squares is lighter?”

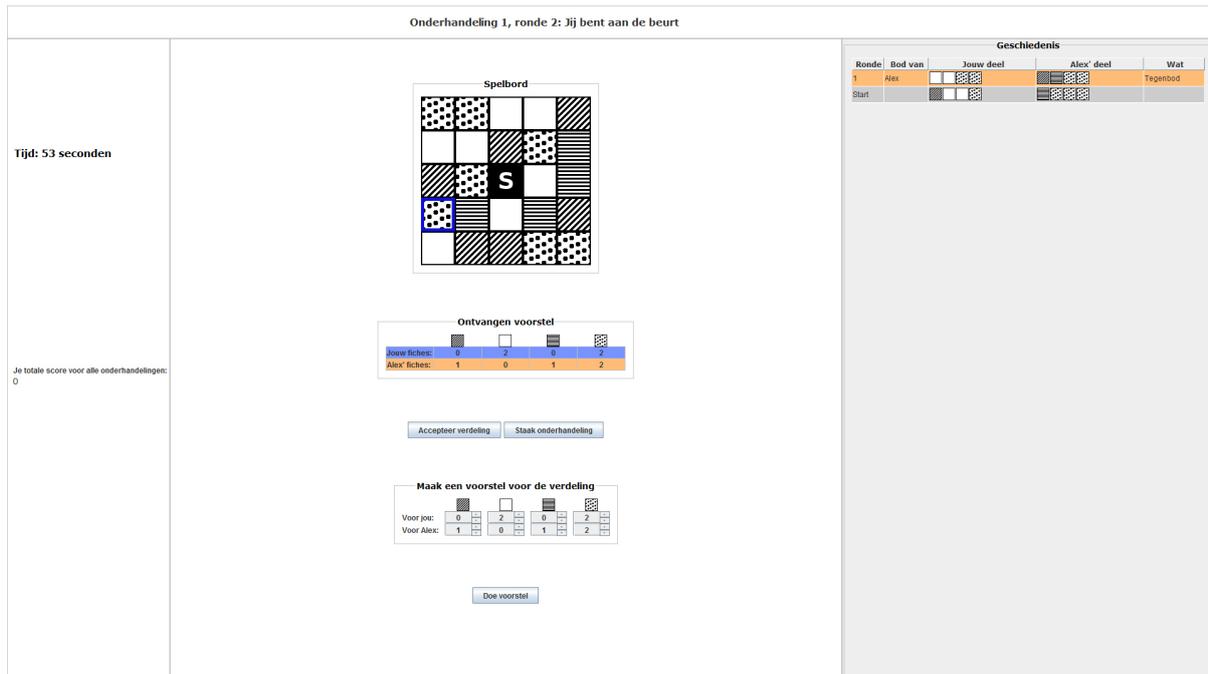


Figure A.3: This is the screen in which colored trails takes place. At the absolute top of the screen, the participant can find the current trial and round of that trial, at the topleft corner the time left is displayed, beneath that is the score at the moment, In the middle of the screen the playing field is displayed. Beneath the playing field the participant can find the offer received from the agent and two buttons to 'accept the offer' or 'stop the negotiation'. Beneath the two buttons an interacteable table is shown, where the participant can make his own offer and click the 'make offer' button below it. Finally at the right of the screen the participant can see the history of his negotiations.

The distribution of the scores per trial of the two groups in Colored Trails

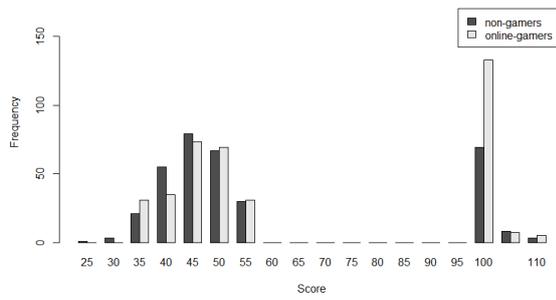


Figure A.4: The dark-gray bars are the scores per trial for non-gamers, the light-gray bars are the scores per trial for online-gamers.

The distribution of the score per trial against a Tom<sub>1</sub> agent in Colored Trails

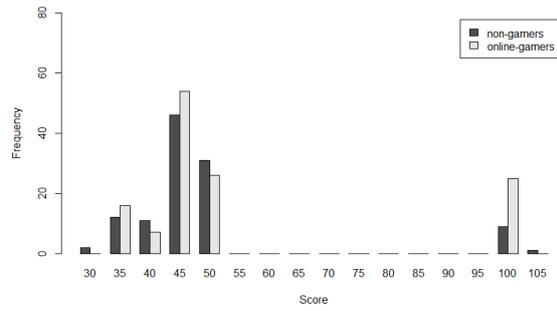


Figure A.6: The dark-gray bars are the scores per trial for non-gamers, the light-gray bars are the scores per trial for online-gamers.

The distribution of the score per trial against a Tom<sub>0</sub> agent in Colored Trails

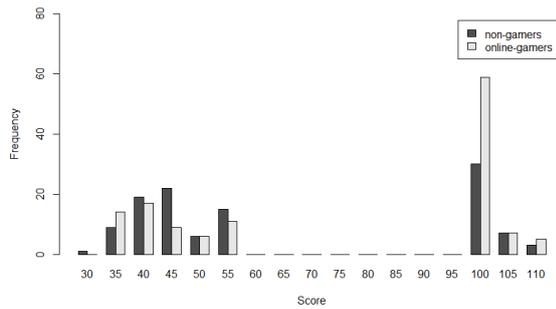


Figure A.5: The dark-gray bars are the scores per trial for non-gamers, the light-gray bars are the scores per trial for online-gamers.

The distribution of the score per trial against a Tom<sub>2</sub> agent in Colored Trails

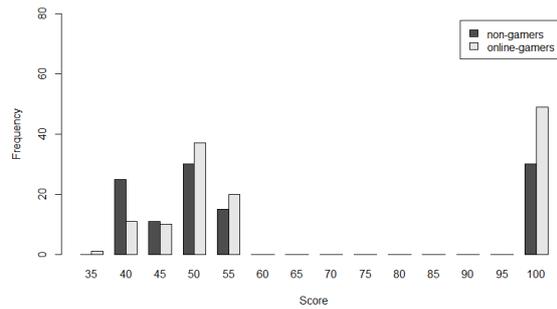


Figure A.7: The dark-gray bars are the scores per trial for non-gamers, the light-gray bars are the scores per trial for online-gamers.